

94272

From: Chan, Christina
Sent: Monday, May 19, 2003 9:10 AM
To: Schnizer, Holly; STIC-Biotech/ChemLib
Subject: RE: RUSH: seq. search in appl. no. 09/647522

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

CRF

-----Original Message-----

From: Schnizer, Holly
Sent: Monday, May 19, 2003 9:00 AM
To: Chan, Christina
Subject: RUSH: seq. search in appl. no. 09/647522

I would like to request a Rush sequence search in the above application which is After-Final.

Please search the commercial and interference databases for SEQ ID NO:5

Thank you.

Holly Schnizer
AU 1653
CM1-9E09
305-3722
mailbox: CM1-9B01

Mary Jane Ruhl
Tech. Info. Specialist, STIC
TC-1600
CM-1, Room 6A-06
Phone: 605-1155

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 5/19/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

```

DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE      Methyl-accepting chemotaxis protein McpX.
GN      McpX.
OS      Rhizobium meliloti (Sinorhizobium meliloti).
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Rhizobiaceae; Sinorhizobium.
OX      NCBI_TaxID=382;
RN      []
RP      SEQUENCE FROM N.A.
RA      Muschler P.F.J., Scharf B., Schmitt R.;
RT      "Methyl-accepting chemotaxis protein McpX of Sinorhizobium meliloti."
RL      Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF312879; AAC34158.1;
DR      HSSP: P02942; I0U7
DR      InterPro: IPR000267; Asp/Glutamase.
DR      InterPro: IPR004010; Cache.
DR      InterPro: IPR004089; Chmtaxis_transd.
DR      InterPro: IPR003660; HAMP.
DR      InterPro: IPR003661; His_kinA.
DR      InterPro: IPR004090; Me_Chemotaxis.
DR      Pfam: PF02743; Cache; 1.
DR      Pfam: PF00672; HAMP; 1.
DR      Pfam: PF00015; MCPsignal; 1.
DR      PRINTS: PR00260; CHEMTRNSDCCR.
DR      SMART: SM00304; HAMP; 2.
DR      SMART: SM00388; HSKA; 1.
DR      SMART: SM00283; MA; 1.
DR      PROSITE: PS00144; ASN_GLN_ASE_1; UNKNOWN.1.
SO      SEQUENCE 788 AA; 8573 MW; 3681BC667DDE83F5 CRC64;

```

```

OY 173 MKQSNLTDSLSIIANPVYKFSNFIQLESRIQSGAFTSLSDAKRAVDLFLYCOL 232
DB 250 LKSVVOOTDKD-SEMRANV-----FISMSKALVNO-----QFAESKRCLEFA---CEI 294
OY 233 VVARETL--VDLAILYKKGNAEHAASAVENANRYNK--ELAADTIDELHLKLIPEQ----- 284
DB 295 LEKKEPALPEVAEAY-----SEVAMQYESMNEFEFAISLQKTLGILEKLPQOQHSEGS 349
OY 285 --ALIG-----AYHPIASSETSKAILNTKYFGVPPVPRPIG 320
DB 350 VSARIGWLLIFSGRVSOAVPYLESAAERLKESEFG-AKHFGVGYNNMG 397

RESULT 13
Q8VZ08 PRELIMINARY; PRT; 617 AA.
AC 08VZ08;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative kinesin light chain protein.
GN ATG31240.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RA Yamada K., Banb J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carlini P., Chen H., Cheuk R.,
RA Hayashitaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neuman G.,
RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At2g31240 (GI:15224637).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY063932; FAL36288.1;
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR: 7.
DR SMART: SM00028; TPR: 8.
SQ SEQUENCE 617 AA; 68744 MW; 589443C876CFC32F CRC64;

Query Match 4.88; Score 112.5; DB 10; Length 617;
Best Local Similarity 22.6%; Pred. No. 15;
Matches 79; Conservative 43; Mismatches 126; Indels 101; Gaps 16;

OY 25 DVNSILTKYETAL--KEASGSNPALEALEGLKEIOTRPRDVGQATKILGVSALGK 81
DB 204 DVSRVLAHVOLELANVNAAGRR--EALENKKSLEITEMTFDESKEGVNRSL-- 258
OY 82 LNSGDATKIISGCLDIAGIATTEGPGVGMIGAVASFEVSSILSLFTGSASKNVAVID 141
DB 259 ---ADAVAV-----LNFNALPYALKALEIHKKEKELGNNAEVAOD 296
OY 142 RAL-----SKRDEAIQRH-----AAGAKRDFAESAFIYQV 172
DB 297 RRLIGVYISGLEQHDKALENRLSORVLRKMGKMLLEIRAEIDIANNAKVALGKYEBAIDI 356
OY 173 MKQSNLTDSLSIIANPVYKFSNFIQLESRIQSGAFTSLSDAKRAVDLFLYCOL 232
DB 357 LKSVVOOTDKD-SEMRANV-----FISMSKALVNO-----QFAESKRCLEFA---CEI 401
OY 233 VVARETL--VDLAILYKKGNAEHAASAVENANRYNK--ELAADTIDELHLKLIPEQ----- 284
DB 402 LEKKEPALPEVAEAY-----SEVAMQYESMNEFEFAISLQKTLGILEKLPQOQHSEGS 456
OY 285 --ALIG-----AYHPIASSETSKAILNTKYFGVPPVPRPIG 320
DB 457 VSARIGWLLIFSGRVSOAVPYLESAAERLKESEFG-AKHFGVGYNNMG 504

```

```

RESULT 14
O99054 PRELIMINARY; PRT; 6713 AA.
AC O99054;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE BbA protein.
GN BBHA OR SA1267.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-21311952; PubMed-11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Iino T., Kanamori M.,
RA Matsunari H., Maruyama A., Murakami H., Hosoyama A., Mizutani U.I.,
RA Takahashi N.K., Kuhara S., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kohara S., Goto S., Yabuuchi J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003133; BAB42527.1;
KW Complete proteome.
SQ SEQUENCE 6713 AA; 722339 MW; AF6EDED226BE4888 CRC64;

```

```

Query Match 4.88; Score 112.5; DB 16; Length 6713;
Best Local Similarity 21.7%; Pred. No. 5,5e+02;
Matches 92; Conservative 55; Mismatches 157; Indels 119; Gaps 17;

OY 16 TSAKHGRSVNSILTKYETALKEASGSN-----EALR-----ALRG----- 53
DB 2037 TEADRDKQIYNDVAVTAATLILDKTAGSNQNKAVAEALQRYNATKRALNGDARLNKKN 2096
OY 54 -LNGEIQTK-----PDRYQATKILGVSQ--SALGKLN-----SGDAT 88
DB 2097 TAKQQLATMGSHTFNAQKANLLEQIERGTIVAGVGIQANAGTLNQAANQLOSTASDKAT 2156
OY 89 KIISGCLDIYAGIATTEGPGVGMIGAV-----ASFVSSILSLFTG--- 129
DB 2157 KSESDYDANADQNAIYNDVATNAEGILISATNNPEAMPDTINQKASQVNSAKSLNGDEK 2216
OY 130 -SSAKNSVAVIDRALSKHRDEAIQRHAGAKRDFAESAFIYQVKKOS----- 177
DB 2217 LAAAKQTKAKDIGHGLDNLN---AQRTAANAEDQANLAAVTAAKKATSLNTRMGNLK 2273
OY 178 -----NLNDS-----LSITANVPYKFSNFIQLESRIQSGAFTSLSDAK 220
DB 2274 HALAEKNDTRRSVYTDADPOKQAYDTAVTQAEATNANGSNANETQVOAALNQLQAK 2313
OY 221 RAVDFILLYCOLVYRETLVDLAILYKKGNAEHA--SAVVENANRYNKLELA--DTLDFL 277
DB 2334 NDLGDNKVAQ--AKTKARALASTYSLNNAQSTASTSIDNLTYYADVTAQANANEL 2390
OY 278 HKLIPQALGAVYHPIASSETSKAILNTKYFGVPPVPRPIGRNRRKFTNSYWNYSIC 337
DB 2391 N-----TANGQLONGINDQNTYKQOVNFI-----DADQ---GKNQAYTNAYTNAGGIL 2435
OY 338 SEA 340
DB 2436 DKA 2438

RESULT 15
O9F5M7 PRELIMINARY; PRT; 788 AA.
AC O9F5M7;

```

RA Wilkinson J.E., O'Bryne C., Shea T., De Loughery C., Caspi R.
 RT "Sequence Analysis of a 101 kb Plasmid Required for Agar degradation
 of a Micrococcia strain."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF393846; AK62869.1;
 DR InterPro: IPR004089; Clntaxis_cransd.
 DR InterPro: IPR000566; Lipoclin_cytfabp.
 DR Pfam: PF00015; MCPsignal.1;
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN.1.
 KW Plasmid.
 SQ SEQUENCE 597 AA; 65969 MW; CA09453850F918B2 CRC64;

Query Match 4.9%; Score 114.5; DB 2; Length 597;
 Best Local Similarity 21.2%; Pred. No. 11;
 Matches 65; Conservative 59; Mismatches 106; Indels 77; Gaps 12;

QY 16 TSARKGRSDVNSLTFRKETALKEASGSENALEA-LEGLK-----EIQKPRV 65
 DB 296 TQVOVSKIDOVSSLEELKASEMRGKVEITTEASVIGKNSENGHLADRINKMEHT 355
 QY 66 GQATRIIGSGALGKLSGATRIISGLIVAGIATTFGPGVGMIGAVASFVSILS 125
 DB 356 AKLSKNSKMSVLSF-KSNEITRIIS-----VIGSIANO-----TNLLA 394
 QY 126 LFTGSSAKNSVAAYIDRLSKHRDEAIORHAAGAKRDEASSAFIOWKQOONLTDLS 185
 DB 395 L-NAATEAQAQSDAGRGFSVADE-IRKLANDSRKRETSFVNEIQOQTN----- 444
 QY 186 IIAANVYKFSNFISGLESNISGAAATTSIDAKRAVDFTLLYCOLVYMETLLVLA 245
 DB 445 -----EANLMSMHGIRGEITTS-----AAMESPOL-----I 474
 QY 246 LYRKNAEHVAAVENANRYNKE-----LAADTLDLFLKLPKALIAVYHPIASSETSK 301
 DB 475 LQSSKDKRKHAEDELSTTGQEMKLVHOVASFNTENIV--VVAEQNSVGTQEIASSAELSS 532
 QY 302 ALLNYTK 308
 DB 533 GMTNYRK 539

RESULT 11
 Q9R9R9 PRELIMINARY; PRT; 306 AA.
 AC Q9R9R9; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 GN FLA.
 OS Aeromonas punctata (Aeromonas caviae).
 OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
 OC Aeromonas.
 OX NCBI_Taxid=648;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SCH3N;
 RX MEDLINE=21295073; PubMed=11401962;
 RA Rabaan A.A., Gryllos I., Tomas J.M., Shaw J.G.;
 RT "Motility and the Polar Flagellum Are Required for Aeromonas caviae
 Infect. Immun. 69:4257-4267(2001).
 DR EMBL: AF198617; AAF19179.1;
 DR InterPro: IPR001492; Flagellin.
 DR InterPro: IPR001029; Flagellin_C.
 DR Pfam: PF00700; Flagellin_C.1;
 DR Pfam: PF00669; Flagellin_N.1;
 DR PRINTS: PR00207; FLAGELLIN.
 DR PRODOM: PD000316; FLAGELLIN.
 SQ SEQUENCE 306 AA; 31771 MW; 22243CB52A1D8757 CRC64;

Query Match 4.9%; Score 113.5; DB 2; Length 306;
 Best Local Similarity 25.3%; Pred. No. 4.5;

Matches 60; Conservative 38; Mismatches 98; Indels 41; Gaps 11;
 QY 13 LAITSARKRSDVNSLTFRKET-ALKEASGSENA-LEALGKGEIOTKPRVQATK 70
 DB 72 ISLAOTAEAGMEVETGMLOMRMTLAQOOSANGSNASKDRALKEVDQLAEINRISTATT 131
 QY 71 IIGSV-----GSALGXKLSNG-DARKIISGCLD-----IVAGIATTFGPGVGMIGAVASFVS 121
 DB 132 FAGTKLLDGSFSGTFQVGADANOTIGFSLAOTGCFSGISGIAKAGATTIDIVSGPAGSVTT 191
 QY 122 S--ILSFTGSSA-----KNSVAVIDRAL-----SKIRD-EAIORHAAGAKRDE 164
 DB 192 ATGISLITFTGSGAGCISTISTOSKAQAVLAADAMLEVDSKRAELGAVQNRDLSTIRNOA 251
 QY 165 ESSAFIOWKQOONLTDLSITIAANVYKFSNFISGLESNISGAAATTSIDAKR 221
 DB 252 NISE--NVSAARSRIODAFATERANM-----TKONILOQAASSTIAQANQ 295

RESULT 12
 Q9SUX2 PRELIMINARY; PRT; 510 AA.
 AC Q9SUX2; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 GN AR2G31240.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Niernhan W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana".
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006593; AAD20669.1;
 DR InterPro: IPR001440; TPR.
 DR Pfam: PF00515; TPR.7.
 DR SMART: SM00028; TPR.3.
 SQ SEQUENCE 510 AA; 56494 MW; 908383937BB8F5F9 CRC64;

Query Match 4.8%; Score 112.5; DB 10; Length 510;
 Best Local Similarity 22.6%; Pred. No. 11;
 Matches 79; Conservative 43; Mismatches 126; Indels 101; Gaps 16;

QY 25 DVNSLTFRKETAL-----KEASGSENALEALGKGEIOTKPRVQATKILGVSGLGK 81
 DB 97 DVREVLAVOLELANVAMNGRRE--EATEINKLSLEIKEMTFDEDSKMGVNRSL- 151
 QY 82 LNSGATRIISGLDVIAGIATTFGPGVGMIGAVASFVSILSFTGSSAKNSVAAYID 141
 DB 152 ---ADAVAV-----LNFEALPYALKALAIHKRELGNNSAEVAVOD 189
 QY 142 RAL-----SKRDEAIORH-----AGAKRDFASSAFIOW 172
 DB 190 RRLIGVIYSGLEQHDRALEONRLSORVLKNGMKLELIRAEIDAANMKVALGKYEBAIDI 249

087527
ID 087527 PRELIMINARY; PRT; 877 AA.
AC 087527;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE ABC transporter ABCA.1.
GN ABCA.1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Anjaard C., Loomis W.F.;
RT "Evolution of the ABC transporter of Dictyostelium";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF491005; AAL99041.1; -
SQ SEQUENCE 877 AA; 97375 MW; DBEC7C7ECCB343D9 CRC64;

Query Match 5.0%; Score 115.5; DB 5; Length 877;
Best Local Similarity 20.9%; Pred. No. 16;
Matches 77; Conservative 52; Mismatches 128; Indels 111; Gaps 17;

QY 7 PWFIV-----LAITSAKH-----GKSDVNSLTKRVEATLKEAGSNEALELEGK 56
DB 507 PWFVLPSTWGLSKVRHIPPYDEDEVRAITKADA-----SNRAPL-IICGLSK 559
QY 57 EIQT--KPDVQCATKIL-----GSVGSALGLKNSGDATK--ISGCLDIYAGIATTPG 106
DB 560 STKLFPRKKYHAKYKLSVSEKGIILGFLSNGCGKTTIGMLTGLEPAGDGLVYG 619
QY 107 GVGKGIGAVASFSVSI-----LSLEFT-----GSSAKNSVAVIDRA-L 144
DB 620 HSVISIAAVRRITSVPOHDILMAEMTAREHLOLFSEKGIPOAERESQIOKVLDOVRL 679
QY 145 SKHREAIORHAAGAKR-----DEAESATLIOWKQOSN-- 178
DB 680 SKISNNLSTYSGMKRRLSVALACIGDKPIIFMDEPTTGVDPSKRHLIDLKSTKNK 739
QY 179 ---LTDSDL-----STIAANVPYKFSNFIGLESRIISOGAATSTLSDAKRAVDFTL 227
DB 740 VILLSHDHVEYELADKIVINNEGVMACNGSLQKSYGEGYSNVIYAKSPESIPAVY 799
QY 228 LYQQLVY---MRETL-----VDLILYRKGAHVAASAVENNRNKEAAD- 272
DB 800 EFTVLSIPCKFMKOSALQNLNFGFVPTIDHQT-----ANFKOLEEITNDPNQIMRDM 854
QY 273 -----TLD 275
DB 855 SVSHSTLD 862

RESULT 9
082B23 PRELIMINARY; PRT; 800 AA.
AC 082B23;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative Rhs accessory genetic element.
GN YP03606.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parhill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Kariyeh A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague";
RL Nature 413:523-527(2001).
DR EMBL: AJ414158; CAC93075.1; -
DR InterPro: IPR001757; ATPase_E1-E2.
DR PROSITE: PS00154; ATPase_E1-E2; UNKNOWN_1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 800 AA; 87222 MW; 003F04A89D99298 CRC64;

Query Match 5.0%; Score 115; DB 16; Length 800;
Best Local Similarity 17.3%; Pred. No. 15;
Matches 101; Conservative 79; Mismatches 202; Indels 202; Gaps 24;

QY 12 VIATSAKHGRSDVNSLTKRVEATLKEA-----SGSNEAL-----EALGKGLQIKP 62
DB 1 MEAHDKANNNKADRTGTPTPHIVSPAPATVSGAQMTEAGGIGDVGTAQR----- 55
QY 63 DRVGATKILGVSALGLKNSGDAT-----KIISGC--LDIYAGIATTPGGPVG 110
DB 56 -RINGAAVAKTALDAKAKVLDGVTPTNIVATGPAPTVSGAQMTEAGGIGDVGTAQR 114
QY 111 MGIGAVASFSVSI--LFTGSSAKNSVAVIDRALSKHDEAIORHAAGAKRDFASSAFT 170
DB 115 KRINQAE-----YAKTLEAKAKYLDGCVTPM--LAAGAGAGLPDSAAATIS 162
QY 171 QVMKQSNLTDLSITIAANVPYKFSNFIGLESRIISOG--AATSTLSDAKRAVDFTL 229
DB 169 RLVKQPSGL---QFLTASLPPQTFVAVDFLSLSEMLSPFVLANGLASADPAVDPAVL 219
QY 230 ---COLVYRE-----TLVDLALTYRG 250
DB 220 DEDATLFTIREKVALQSRITGVAVSFEQDGTGFHQRYSNVPALMRTSLRNNAIFQQA 279
QY 251 NAEHVASAVENNRNRYNK-----ELADTIDFLAKLIPLEAL----- 286
DB 280 SYEITITTLKENGINDFAFGFRHHPVAREFCVOQESDFDIQRLTAEGFYFEFSA 339
QY 287 -----IGAV-----YHP-ISASEISKALINTKTYGVPDVPPIGRKRTKFTN 328
DB 340 GKNTVYADVDGSLPKGASLPYNPVAAQAOELSTTFTRSAQVR--PAMVOLKDYTFKN 397
QY 329 STY-NTYSICS-----YANVGNVFRGCSNVRNINIVSKMSDF 367
DB 398 PMAAFAFSQSGELONQRPDYEHDPGFRKDAQHODFTRLDLRLNDANIGQASND 457
QY 368 YTMNSDRRKLKTYTKHDG--GFW-----GLDEDPGDOG-----HMRIFPL 407
DB 458 FTLPQGLFSIL--NHPRDLNHNQDLGTOHSGKMOALBQASGDQGVTLNHFIFP- 514
QY 408 RHGKYVSSKRWPNFMFMSSASGYSRWENN-----GPQ 444
DB 515 -----HTQWTRPTPLAKPMDGPQ 533

RESULT 10
093P84 PRELIMINARY; PRT; 597 AA.
AC 093P84;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MS147, putative chemotaxis protein.
OS Microscilla sp. PREL.
OC Plasmid PSD15.
OC Bacteria; CF8 group; Flexibacter group; Microscilla.
OX NCBI_TaxID=155537;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PREL;
RA Zhong Z., Toudkarian A., Helinski D., Knauf V., Sykes S.,

OY 300 SKAIIYTYFGVDPVPRIGN--RRYKFTNSYMTYSICSEAYMGNYMFGSGSNRPN 357
 Db 510 ADKLLYTDY---PEEKTLITLNRKHA-----RHICKGIARPKLV---Y 551
 OY 358 IYVSKMSD-----GFTYMSDRRKLYITKHDQW 387
 Db 552 YRAKFIYVNNYKGRYNEEDTKLKAHSHLGNMW 586

RESULT 6

O9XOM7 PRELIMINARY: PRT: 530 AA.
 AC O9XOM7,
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Methyl-accepting chemotaxis protein.
 GN TM1143.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB / DSM 3109;
 RX MEDLINE=9287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima.
 RL Nature 399:323-329(1999).
 DR EMBL: AE001771; AAD36219.1;
 DR HSSP: P02944; 1QV7.
 DR TIGR: TM1143;
 DR InterPro: IPR004089; Chmtaxis_transd.
 DR InterPro: IPR003660; HAM.
 DR InterPro: IPR004090; Me_chemotaxis.
 DR InterPro: IPR000727; T-SNARE.
 DR Pfam: PF00672; HAM; 1.
 DR Pfam: PF00015; MCPsignal; 1.
 DR PRINTS: PR00260; CHEMTNSDCCR.
 DR SMART: SM00283; MA; 1.
 DR Complete proteome.
 SQ SEQUENCE 530 AA; 57928 MW; C04149A4F6890CE CRC64;

Query Match Best Local Similarity 5.0%; Score 116; DB 16; Length 530;
 Matches 54; Conservative 63; Mismatches 121; Indels 40; Gaps 9;

OY 15 ITSAGHGRSDVNSLTFRKET---ALKEASGSENALELBEGIKETQTPRPGVQATRI 71
 Db 248 IINOLLGISKENDINSTRLESISASVQETTAGSEETSKTKNIASQAQAFADOSTQL 307
 OY 72 LGSVSAIGKINSGDATKIISGCDIYAGIATPGPVGMGAVASFPSSLSLFTGSS 131
 Db 308 AKKADALAKVY--EVTNIMSASAKDVERVSEF---OKGAETTSFVETTNAL----- 356
 OY 132 AKNSVAAYIDRALSKHRDEAIORHAAGAKRDFEASSAFIOWAKQOSNTDPSLTIANV 191
 Db 357 AEQNLALALMAIEAAR-----AGEAGRGFAVVADEIRKLAESQASQASEVNRVVEI 409
 OY 192 -----PYVKFSNFIIGLESISOGAATISLSDAK-----RAVDFILLYCOLVMRET 238
 Db 410 RSIADNACKVSS---ETIARVEBGKLADEADEKINSIVGAVERNEMQINAAIEBOT 466
 OY 239 LLYD---LAILYRKNAEHAASAVENANRVKKELAADT 273
 Db 467 AADEITTAATENAKNAEETINSYKVENARLOEISAST 504

RESULT 7

O92B35 PRELIMINARY: PRT: 1571 AA.
 AC O92B35;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein lin1716.
 GN LIN1716.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Checouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Doman E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Emtian K.-D., Esibi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkut G.,
 RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Rammel B., Rose M., Schlueter T., Simoes N., Thierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
 RT Comparative genomics of Listeria species.
 RL Science 294:849-852(2001).
 DR EMBL: AL596169; CAC96947.1;
 DR Listlist; LIN1716;
 DR InterPro: IPR001525; C5_DNA_meth.
 DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37; 1.
 DR PROSITE: PS00095; C5_MTASE_2; UNKNOWN.1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1571 AA; 167910 MW; B3BE9809D9581P6 CRC64;

Query Match Best Local Similarity 5.0%; Score 116; DB 16; Length 1571;
 Matches 84; Conservative 74; Mismatches 148; Indels 108; Gaps 18;

OY 11 IYLAITSAGHGRSDVNSLTFRKETALKEASGSENALELBEGIKETQTPRPGV--QA 68
 Db 141 LVVAGSKAMFNASDINELNVDAFKNA-----BSVNRWSTTIDNIGLAOG 190
 OY 63 TKI-----LGSVSAIGKINSGDATKIISGCDIYAGIATPGPVGMGIGAVSPV--- 120
 Db 191 TALDLAATYGDWSTSMG--LSTQEARKMSYWDLA-----GMLASFKNID 234
 OY 121 -----SLSLFTG--SSAKNSVAAYIDR-----ALSKHRDEAIORHAAGAKRDFEASSA 168
 Db 235 IDRANTALNVGVTGETELKSLGTYMTQTNLEQFALFEGAGKVAKSSTEVTKONLARE-- 292
 OY 169 FIOVWKOOSNTDPSLTIANVPPYKFSNFIIGLESISOGAATISLSDAKRAVDLIL 228
 Db 293 -----KAQKLINEAKKEGKNSLAREQONKLOETQARTSBC--AKVNLKDKKO--DELYR 344
 OY 229 YCOLVMRETLVLDAILYRKNAEHAASAVENANRV-----NKELAADTDLDFKLHLEQ 284
 Db 345 LRYNYVMKOT-----TNAHGDFKNTSDQANNSRPFSESVKELASNAQOFLPLIT--T 395
 OY 285 ALIAYVHPIASSETSKAIIYTYFGVDPVPRIGNRRKFTNSYMTYSICSEAYM-- 342
 Db 396 PLI-----IQAOSFTKRL-----SDIPGAVKMKKEKFPAP--EVEFETVGDFFKND 439
 OY 343 -----GNYMFGSGSNRPNIRRSKMSDGFYTMENDRRK 377
 Db 440 LIPSARELAKSMGPGFTIEGVLAFKGLGIVLNTTVIPAFRAVTKFTRENDPSMK 493

RESULT 8

Db	Accession	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	286 LIGA-VYHPISASTSKAILN 305	5.18	506	66	39	99	115	
Db	486 QVSAMVKHRASRLIEDAVTLH 506							
RESULT 4								
ID	0823M0	PRELIMINARY;	PRT;	506 AA.				
AC	0823M0;							
DT	01-MAR-2002 (TrEMBLrel. 20, Created)							
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)							
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)							
DE	Aerolotaxis receptor protein.							
GN	stry3395.							
OS	Salmonella typh.							
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;							
OC	Salmonella.							
OX	NCBI_Taxid=601;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=CT18;							
RX	MEDLINE=21534947; PubMed=11677608;							
RA	Parthill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.							
RA	Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebalha M.,							
RA	Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,							
RA	Croun A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,							
RA	Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,							
RA	Krogh A., Larsen T.S., Leather S., Moule S., O'Gara P., Parry C.,							
RA	Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,							
RA	Whitehead S., Barrrell B.G.;							
RT	"Complete genome sequence of a multiple drug resistant Salmonella							
RT	enterica serovar typh CT18.";							
RL	Nature 413:848-852(2001).							
DR	EMBL; AL627278; CAD07741.1;							
DR	InterPro: IPR004089; Chmtaxis_transd.							
DR	InterPro: IPR003660; HAMP.							
DR	InterPro: IPR001610; PAC.							
DR	InterPro: IPR000114; PAS_domain.							
DR	Pfam; PF00672; HAMP.1.							
DR	Pfam; PF00015; MCPsignal; 1.							
DR	Pfam; PF00785; PAC; 1.							
DR	SMART; SM00283; PAC; 1.							
DR	SMART; SM00086; PAC; 1.							
DR	TIGRFAMS; TIGR00229; sensory_box; 1.							
DR	TIGRFAMS; TIGR00229; sensory_box; 1.							
KW	Receptor; Complete proteome.							
SO	SEQUENCE 506 AA; 55139 MW; DED2387AB045FB30 CRC64;							
QY	Query March	5.18	Score 118.5;	DB 16;	Length 506;			
Db	Best Local Similarity 21.22; Pred. No. 4.3;							
Db	Matches 66; Conservative 39; Mismatches 99; Indels 115; Gaps							
QY	24 SDVNSLTKVTALKEASGNEALELEGLKEIOTRPPRVGQATKILISVGSALCKLN 83							
Db	262 SQVSSVRNGSRRLKGNNDLNEHTRQRYENVQ--ETVYTNQMAESVTLNSETSAADKLS 320							
QY	84 -----SSDARKIISGCLDIYA-----GIATTF-----GG 107							
Db	321 MAASAAATGGEADMDYVTKTDDIAHSTORIGITTTLLINDIAFOPTNIALNAVEAARAG 380							
QY	108 PYGMCIGAVASFYVSIISLFTGSSAKNSVAVIDRALSKHDEAIQRMAGAKRDFESS 167							
Db	381 EOGKGFAVVAEYVHLAS--KSAANAADIRKLIDASATKYQSSGEVYHAAGRITD----- 433							
QY	168 AFIOVMKQOSNLWTDSDISIAANVPYKFSNFIGOLESRISQGA--ATFSLSDAKRAVD 225							
Db	434 ---DIVAQOVAVT-----LLIARIISQSQEOETDGLISLITRAYD----- 468							
QY	226 ILIVCOLVYMKETLLVDALILYRKGNAEHVASAVENMNRVAKELTAADTLDFIHLKLEPQ 285							
Db	469 -----ELNRITQKNA-----LVESA 485							

QY	286	LIGA-VYHPISASETSKALIN	305
Db	486	QVSAWVKHRASRLDVAITVLH	506
RESULT 5			
ID	062187	PRELIMINARY;	PRT; 833 AA.
AC	062187;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Transcription factor.		
GN	TTF1 OR TTF-1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EHRLICH ASCITES;		
RC	MEDLINE=95237206; PubMed=7720715;		
RA	Evers R., Grummt I.,		
RT	"Molecular evolution of mammalian ribosomal gene terminator sequences		
RT	and the transcription termination factor TTF-1,"		
RL	Proc. Natl. Acad. Sci. U.S.A. 92:5627-5631(1995).		
CC	-1- SUBCELLULAR LOCATION: NUCLEUS (BY SIMILARITY).		
CC	-1- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.		
DR	EMBL; X83974, CAA58808.1; -		
DR	MGI; MGI:105044; Tefl.		
DR	InterPro: IPR001005; MYB DNA-binding.		
DR	Pfam; PF00249; myb DNA-binding; 1.		
DR	SMART; SM00395; SANT; 2.		
DR	PROSITE; PS50090; MYB_3; 1.		
KW	DNA-binding; Nuclear protein.		
SO	SEQUENCE 833 AA; 94534 MW; DEB9C8DC6D3ECC6 CRC64;		
Query Match	5.08; Score 117; DB 11; Length 833;		
Best Local Similarity	22.08; Pred. No. 12;		
Matches 100; Conservative	54; Mismatches 127; Indels 174; Gaps		
QY	4 KHL-----PWLEIVLAITSAKHGKSDVNSLLTKETLKAASG-SNEAALEALEGLK---	55	
Db	235 KHLQKVAWP-----DVVOGSPPEISILPPSPPLSSEDLDEKSTAAVFCRSLKKNV	286	
QY	56 ---GEIQTKPRVGOATKIIIGSVSALGKINSGDATKIISGLDIYAGIATTFGPGVGMG	112	
Db	287 FRQSELEPIPD-----SLDDSETISERLD-----STHHGAVAGAG	321	
QY	113 -----IGAVASFPVSSILSLFTGSSAKNSVAAYIDR-----	142	
Db	322 ECSESTKSHSTIKKSKKKKHHVSVALVTSOSASVATDSKAN---ALVDSESGGAVREED	378	
QY	143 -----ALSKRDEAIDR-----IAAGAKRPFASASAFIYMKOOSMLTSDLS	185	
Db	379 VDRHRAEAEQAQSTEHNR-EAMQRLEPTHEESNNSASNAARIHSEDRRSSDDSDVD	437	
QY	186 IIAANVPYKESNFTIGOLESRISGGAATTSLSDAKRAVDIFILLYCOLVMRETFLYDLAI	245	
Db	438 LGSQ---VROQREFTPIQIER---AAAT-----IRRYRQDLG-469		
QY	246 LYRGNAEHVA-----SAVENANRVKELAAIDLFLHLKLPDQALIGAVYPIASST	299	
Db	470 LFEKPKAQGAIRGKRSKAK---NKQLEKNVQDPL-----SLTQIES	509	

OY 138 AYDR--ALSKHDEAIOHRAAGKRDFAESSAFIOVMKQSNLTDSDLS--TIANVPEY 193
 DB 414 AVGEGAVNALGKHIGSSIRKAGADYTNLNGES-----YPELDEDIADOTATDFEN 464
 OY 194 YKFSNFIQLESIRISGAATTSISD-----AKRAVDFILLYC-----OLVAKRETLVDL 243
 DB 465 WADNPNSAVASRLQAFPNASDAELAKKRAVATNSDQSRFNTQKRODQNNKKT----- 519
 OY 244 ALLYKRGNAEHVASAVENANRVNKEADTLDFELKHLIPEQALICAVHPISASETSK-- 301
 DB 520 ATPY--GNARDLVNAATNA--FOKGECDKDTFMSQLPEN-----MSAEKEKOW 566
 OY 302 -----AIIIN-----TKYFGVP--DVPRPIGNRRYKFTSNYNTYSICEAIVMGNM 346
 DB 567 NDHLNAKYOGFENHAEQATKAGAMPVADAKQKQNN--LFQKSYNKKDPAQSOLALG-- 621
 OY 347 FRGGSNVANPNIRVSKMDGF--TYMENSDR-----RKLYTKHQDQWG 388
 DB 622 -----KYGKAEGYKGETLESQHONMGVGLVGAATAFKKGYTADHKAGF- 668
 OY 389 WGLDEDPGDOGHMRFILRHGKYVVSCKRMPNFMVMESSASGY 433
 DB 669 ---WKQFPAD-----MSQOEKAAMKHLIDQKVGQF 696

RESULT 2

ID 08ZG07 PRELIMINARY; PRT; 595 AA.
 AC 08ZG07;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE Putative methyl-accepting chemotaxis protein.
 GN Yersinia pestis.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia; Yersinia.
 NCBI_Taxid=632;
 RN 11)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Churcher C., Mungall K.L.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Fellwell T., Hamlin N., Hollroyd S., Jagsels K., Karlyshev A.V.,
 RA Leather S., Mouton S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.,
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RL Nature 413:523-527(2001).
 DR EMBL; AJ414147; CAC90064.1;
 DR InterPro: IPR004010; Cache.
 DR InterPro: IPR004089; Chmtaxis_transd.
 DR InterPro: IPR003660; HAMF.
 DR Pfam; PF02743; Cache; 1.
 DR Pfam; PF00672; HAMF; 1.
 DR Pfam; PF00015; MCPsignal; 1.
 DR SMART; SM00304; HAMF; 1.
 DR SMART; SM00283; MA; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 595 AA; 61787 MW; 6D8962C87F7B2E5A CRC64;

Query Match 5.1%; Score 119.5; DB 16; Length 595;
 Best Local Similarity 23.8%; Pred. No. 4.7;

Matches 57; Conservative 39; Mismatches 103; Indels 41; Gaps 11;

OY 21 GKRDVNSLTKEVETAKKESG-----SNEALEALEGLK--GEIOTKPDVRCQ 67
 DB 371 GRRESAASLQOTSAALEJOISATVAQASASARQANNVAFSSEASAGDVITK----- 424
 OY 68 ATRIGSVGSALGKINGDTRKTIISG-----CIDIYAGIATTTGGGVGKIGAVASVFS 121

DB 425 VITMESIEKASGKI--GDITSVIDGIAFOQINILALNAVEARAGEGRGAVAGEV- 481
 OY 122 SLISLFTGSSAKNSVAVIDRALSKHREDAID--RHAAGAKRDPAES--SAFIOVMKQSNL 179
 DB 482 RILAOBSAQAAK--EIKALIESVSVASQSSQVROASNMATDIVASVDVITMSEITNA 540
 OY 180 TDSDSLIIAANVPYKFSNFISQIGLESRIISQGAATTSLSAKRAVDFILLYCOLVMMRETL 239
 DB 541 ADEQMR-----GHEHNSAVAQDPTMVOQNALVQESTNASSA---LQAOADLTDTV 590

RESULT 3

ID 08ZLX8 PRELIMINARY; PRT; 506 AA.
 AC 08ZLX8;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE Aerotaxis sensor receptor, senses cellular redox state or proton
 DE motive force.
 GN AER OR STM3217.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 NCBI_Taxid=602;
 RN 11)
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2."
 RL Nature 413:852-856(2001).
 DR EMBL; AE008848; AL22090.1;
 DR InterPro: IPR004089; Chmtaxis_transd.
 DR InterPro: IPR003660; HAMF.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam; PF00672; HAMF; 1.
 DR Pfam; PF00015; MCPsignal; 1.
 DR Pfam; PF00785; PAC; 1.
 DR SMART; SM00283; MA; 1.
 DR SMART; SM00086; PAC; 1.
 DR TIGRFAMS; TIGR00229; sensory_box; 1.
 DR Receptor; Complete proteome.
 SO SEQUENCE 506 AA; 55063 MW; 70C37C6E7C1BD150 CRC64;

Query Match 5.1%; Score 118.5; DB 16; Length 506;
 Best Local Similarity 21.2%; Pred. No. 4.3;

Matches 68; Conservative 39; Mismatches 99; Indels 115; Gaps 11;

OY 24 SDVNSLTKEVETAKKESGNEALEALEGLKGEIOTKPDVRCQATKLGVSGLGKIN 83
 DB 262 SOVSVRNGSRILAKGNNDLEHRTQYVENVO--ETVTTNMASVSKINSETASADKLS 320
 OY 84 -----SGDAKRIISGCDIYA-----GIATTF-----GG 107
 DB 321 MAASSAATOGGEANDVYIKTMDIDIAHSTQRTITTLINDIAFOQTNILALNAVEARAG 380
 OY 108 PYVGIGIAVASVFSIISLTGSSAKNSVAVIDRALSKHREDAIOHRAAGAKRDPFAESS 167
 DB 381 EOGKGFVAVAGEVHILAS--RSANAAANDIRKLIDASATKYQSGSEGVHAAGRITMD----- 433
 OY 168 AFTQVMKQSNLTDSDLSIIAANVPYKFSNFISQIGLESRIISQGA--ATISLSDAKRAVD 225
 DB 434 ---DIVAOVONV-----LILARISSQSTOEOTDLSILTRAVD- 468
 OY 226 ILLXCOLVMMRETLVDLALILYKRGNAEHVASAVENANRVNKEADTLDFELKHLIPEQA 285

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 11:14:48 ; Search time 89 Seconds

(without alignments)
1041.811 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322
Sequence: 1 MILKHPFLFVLAITSAKH.....SGYIRSWENPQGHMSIT 450

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPTREMBL.21.*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.todent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137	5.9	1222	2	Q9X349 bacillus an
2	119.5	5.1	595	16	Q8ZGQ7 yersinia pe
3	118.5	5.1	506	16	Q8ZIX8 salmonella
4	118.5	5.1	506	16	Q8ZIM0 salmonella
5	117	5.0	833	11	Q6Z187 mus musculu
6	116	5.0	530	16	Q9XOM7 thermotoga
7	116	5.0	1571	16	Q9ZB35 listeria in
8	115.5	5.0	877	5	Q8T5Z7 dictyostell
9	115	5.0	800	16	Q8B23 yersinia pe
10	114.5	4.9	597	2	Q93P84 microscilla
11	113.5	4.9	306	2	Q9R9R9 aeromonas p
12	112.5	4.8	510	10	Q9SUX2 aradionopsis
13	112.5	4.8	617	10	Q8VZ08 arabidopsis
14	112.5	4.8	6713	16	Q9JUS4 staphylococ
15	112	4.8	788	2	Q9F5M7 rhizobium m
16	112	4.8	788	16	Q9ZSH9 rhizobium m

17	112	4.8	2026	2	Q9JPJ4 neisseria m
18	111.5	4.8	655	5	Q94479 dictyostell
19	111.5	4.8	656	16	Q9WYN0 thermotoga
20	111	4.8	865	2	Q9XDU1 clostridium
21	111	4.8	1215	16	Q8XJ10
22	111	4.8	1657	5	Q9BJ59 leptomonas
23	110.5	4.8	656	16	Q9WYR0 thermotoga
24	110.5	4.8	661	16	Q9X020
25	110.5	4.8	1061	2	Q8RNY3 halobacteri
26	110	4.7	810	17	P82857
27	110	4.7	2297	16	Q8YI08
28	109.5	4.7	874	2	Q06277
29	109.5	4.7	6713	16	Q931R6
30	109	4.7	541	16	Q9HUP8
31	109	4.7	612	16	Q9RX1
32	108.5	4.7	456	2	Q93UJ9
33	108.5	4.7	659	16	Q9KVP3
34	108	4.7	492	16	Q8UHY4
35	107.5	4.6	387	2	Q86266
36	107.5	4.6	2155	16	Q8UPP9
37	107	4.6	482	11	Q9Z3V9
38	106.5	4.6	692	16	Q8U1B0
39	106.5	4.6	1098	16	Q50733
40	106.5	4.6	3183	2	Q8RR93
41	106	4.6	1975	16	Q9K0S7
42	105.5	4.5	387	2	Q86265
43	105.5	4.5	667	16	Q8XRM8
44	105.5	4.5	2504	2	Q85160
45	105	4.5	384	2	Q05492

ALIGNMENTS

RESULT 1

ID Q9X349 PRELIMINARY: PRT: 1222 AA.

AC Q9X349: 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PX01-79.

OS Bacillus anthracis.

OC Bacillus anthracis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI_TaxID=1392;

RP [1]

RC MEDLINE-99445483; PubMed-10515943;

RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,

RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,

RA Martinez Y., Rieke D., Svensson R., Jackson P.J.;

RT "Sequence and organization of px01, the large Bacillus anthracis

RT plasmid harboring the anthrax toxin genes."

RL J. Bacteriol. 181:6509-6515(1999).

DR EMBL: AF065404; AAD32383.1; --

SO SEQUENCE 1222 AA; 130407 MW; A7C117874D097E5B CRC64;

Query Match Score 137; DB 2; Length 1222;

Best Local Similarity 23.0%; Pred. No. 0.83; Mismatches 173; Indels 130; Gaps 23;

Matches 107; Conservative 55;

21 GKRSDV-NSLTFKVFETALKEASGSNEALEGLKEGIOTKPDVGAATKIL--GSVGS 77

DB 310 GRRDVRNRSKVRPDDGDKER--TKEDADNPKDALGANIGTDIGTSRRARMLKAGOVGS 367

QY 78 ALGKLNSGDATKTIISGCLDIIVAGIATTFGGPVGMGIGAVASFSLSLFTGSANNSVA 137

DB 368 ALGKATGTG-----LAGMAA--GGALGPGMVGAVSGVSGALGAPGAAGRSVA 413


```

Db 714 EYLTKIESIPERKVAACVHAGGSGKSHAIQALREIGSGITVVLPTNELRLMWSK 773
QY 100 GIATP-----FGG-----PVGMGIGAVASFASSI-LSLFTGSSAK 133
Db 774 KVNTEPEYMERTEKALIGGTGIVITDDYSKLPGR-IALVSFSKIKILITIGTSRQ 832
QY 134 NSVAAYIDRLSKHREDAIQ-----RHAAGA-----KRDAESSAFTQVKQOONLTDSD 183
Db 833 SVYHETSDDASIRHIGPATEVFAKCYRYLNATHRNKDLAN--MLGVYSERTGTETIS 889
QY 184 LSI-----IAANP-----YKKSNTFGOLESR-----SGGA 212
Db 890 MSEEFLGVEPTLVPSDEKRLVMTGTRNDFTTYAGCGGLPKPKVQIYLDHNTOVCSANVM 949
QY 213 TTSLSDAKRAVDFT-----LLYCOLVARETLVLVDLALYRKGAHEV 255
Db 950 YTLASRATDRHINTNSANSASFWEKLDSPYLTPLTSVREQALRE---YEPVEAPEI 1005
QY 256 AS-----AVENANRYNKEFLADTLD-----FLHKLIP 282
Db 1006 REPEPQTHMCVENEESELEEKELLEDREIHEAHGSHNCVQEDTIVQLFSHQAK 1065
QY 283 EQALIGAVYHPISASESKA-----ILNTYKYGVDVPPPTGNRRKFT 327
Db 1066 DETLLAATTDARKLTSNQSNSNFRFYSKRDIGVLEPLNTQKANGLPKEPIP-----FS 1118
QY 328 NSYMTNTYISCEAYMGNMGRCGSNVNPNIRVSKMSDGFYTMENSDRRKLYITKHDGM 387
Db 1119 QEVWEA---CAHEVQSKYLSKSCNKLINGVROS-----PDDEKNIMVFLKSQ-- 1164
QY 388 GWCITLDEDDGDCGHMFIP-LRHGKYV-----SKRWPMFMVMESSASGYR 435
Db 1165 -WTKYKKEKG-----LPRKPGQTIAAFYQQTVMFGTMARYMFMFQAFQPKVEFI- 1215
QY 436 SWENNPPOGHNSIT 450
Db 1216 NCETTPEDMSANAALS 1230

RESULT 14
ID PYRB_ARCFU STANDARD: PRT: 299 AA.
AC 030130:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aspartate carbamoyltransferase (EC 2.1.3.2) (Aspartate
  transcarbamylase) (ATCase).
GN PYRB OR AF0106.
OS Archaeoglobus fulgidus.
OC Archaea: Euryarchaeota: Archaeoglobi: Archaeoglobales:
OC Archaeoglobaceae: Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=96049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
  Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
  Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
  Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
  Kinknes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
  Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,
  Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
  Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
  Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
  Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
  reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -I- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate -> phosphate

```

```

CC + N-carbamoyl-L-aspartate.
CC -I- PATHWAY: Pyrimidine biosynthesis; second step.
CC -I- SUBUNIT: Pyrimidine biosynthesis; second step.
CC -I- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001099; AAB91125.1;
CC HSP: P00479; 3CSU.
CC TIGR: AF0106;
CC InterPro: IPR002029; Asp/Ori-Cotranf.
CC InterPro: IPR002082; Asp-CarbMlttransf.
CC Pfam: PF00185; OTCace; 1.
CC Pfam: PF02729; OTCace_N; 1.
CC PRINTS: PR00100; AORCASE.
CC TIGRPFAM: TIGR00670; asp_carb_tr; 1.
CC PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
CC Pyrimidine biosynthesis; Transferrase; Complete proteome.
CC SEQUENCE 299 AA; 33633 MW; 70F2FC2EBACDD720 CRC64;

Query Match 4.4%; Score 101.5; DB 1; Length 299;
Best Local Similarity 22.6%; Pred. No. 4.6;
Matches 75; Conservative 38; Mismatches 122; Indels 97; Gaps 16;

QY 23 RSDVNSILTKVETALKEAGSNEALDEALG-LKGEIOTKPPRVGQATILSGSALGK 81
Db 13 REDINVLAKKEFEEDVANG--EKKRILEGKILGLNLFEP-----STTRSSFETAMR 65
QY 82 LNSGDAATKIIISGCLDIAGIATTFGGPVGMGIGAVASFSYLSLFTGSSANNSVAAYID 141
Db 66 L-----GADV-----INMTAQAESSI-----ANGETLADTI 91
QY 142 RALSKRDEAIIORHAAGAKRDEASAFIQVK-----QSNLSDSLSTIANAPYVK 195
Db 92 RYVSGTCDAIVTRHPEGAARPAENSVPVINAGDAGQHPOTYLLD-----YT 142
QY 196 FSNFTIGOLESRISOGAATTSLSDAK--RAVDFTLLYCOLVYVR-----ETLV 241
Db 143 IKKEGGRD-----GITLIMGDLKYSRTIHSILKALFLDMRIVLISPEALALPDITL 197
QY 242 DLAIILYRKGNAMHVASVNE--NANVYNKELADTIDFLHKL-----IPQALIGA----- 289
Db 198 DVSAEIRRARLEEVISEIDVLYVTRIQKREFPEDEEY--RKVSGSYRITAEITLSAKDSMI 256
QY 290 VYHPI-----SASETSKALINTYKFGVP 313
Db 257 VMHPLPRVDEIHPSVDSITKHARYFOOSFYGVF 288

RESULT 15
ID FLAA_PSEAE STANDARD: PRT: 393 AA.
AC P21184; 030388; 033928; 068382;
DT 01-MAY-1991 (Rel. 18, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE A-type flagellin.
GN FLIC OR FLAA.
OS Pseudomonas aeruginosa.
OC Bacteria: Proteobacteria: gamma subdivision: Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N. A., AND SEQUENCE OF 1-5.
RC STRAIN=PAK;
RX MEDLINE=91072275; PubMed=2123866;
RA Totten P.A., Lory S.;

```

```

Db 410 LIDASADKVGSGSOQVHAAGRTME-----DIYAQYKNVT----- 443
QY 199 FIGUESISISGA--ATSLSDAKRAVDPIILYCO--LVARETLLVDLALITKRGNAEH 254
Db 444 ---OLIAOISHSTLEADGLSITRAVDLNLITOKNAELVEESAOVSAMV-----KH 493
QY 255 VASAVENA 262
Db 494 RASRLEDA 501

RESULT 12
RRPO_PVXCP STANDARD: PRT: 1456 AA.
ID RRPO_PVXCP STANDARD: PRT: 1456 AA.
AC P22591;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA replication protein (165 kDa protein) (ORF 1) [Contains: RNA-
directed RNA polymerase (EC 2.7.7.48); Probable helicase].
OS Potato virus X (strain CP) (PVX).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=12184;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90364772; PubMed=2392880;
RA Orman B.E., Celnik R.M., Mandel A.M., Torres H.N., Mentaberry A.N.;
RT "Complete cDNA sequence of a South American isolate of potato virus
X."
RL Virus Res. 16:293-306(1990).
CC -1- FUNCTION: RNA-replication. The central part of this protein
CC possibly functions as an ATP-binding helicase.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M31541; AAA47178.1; -
DR EMBL: X55802; CAA39324.1; -
DR PIR: S14005; S14005.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
KW Transferase.
FT NP_BIND 735 742 POTENTIAL.
SQ SEQUENCE 1456 AA; 165301 MW; 489BA57EA070BD2E CRC64;

Query Match 4.4%; Score 103; DB 1; Length 1456;
Best Local Similarity 18.7%; Pred. No. 28;
Matches 104; Conservative 70; Mismatches 187; Indels 194; Gaps 26;

QY 52 EGLKGELQTRDVRGQATKILGVS-----ALGKLNSG-DATKIISG---CLDIVA 99
Db 714 EYLNGKISLPERKVAACVHAGSGSGSHAIKALRELGSGDITVLPLTEALDMSK 773
QY 100 GIATT-----FGG-----PYMGIGAVASVSSI--LSLFTGSSAK 133
Db 774 KVPNTEPYMFTYERKALIGTGSIIVFDYSKLPGY--IEALVSTFKIKLITLIGDSRO 832
QY 134 NSVAVIDRALSKHDEAIO-----RHAAG-----KDFESSAFIQVMKOOSLUTSD 183
Db 833 SVYHETSDASTIRLGPATEVPAKCRYLNTATHRNKKDLAN--MLGVSTERTGTTEIS 889
QY 184 LSI-----IAANVP-----VYKFSNFIGLESRI-----SOGAA 212

```

```

Db 890 MSSEFLGCVPLTVPSEDKRRLYMGNGRNDFTYAGCGELTKPKYQIVLDHNTQVCANVM 949
QY 213 TTSISDAKRAVDFT-----LYICQLYMRETLVDLALITKRGNAHV 255
Db 950 YSALSRAVDRIHFTINTSANSAPFEKLDSTPYLTKFTLSVREHALKE---YEPAAEAPI 1005
QY 256 AS-----AVENANRVKNELAADLTID-----FLHKLIP 282
Db 1006 KEPEPQTHMCYENESVLEETKEELKEFDEIHSDAHGNSCQVTEPTTQILSHQBAK 1065
QY 283 EQALIGAVYHPIISASETSA-----ILNRYTFGVDPVPRIGNRKRYKT 327
Db 1066 DETLLMATIDARLKTNSQESNPREFLSKRDIGVLFUNYQAMGLPKPEIP-----FS 1118
QY 328 NSYNTYSIGCEAYMGNTMGNGCCSNVRNPNTRVSKMSGFTMTENSDRKLYITKHQGM 387
Db 1119 QEVWEA---CAHEVQSKLSKSCNCLNGTVROS-----PDFENKIMVFLKSO-- 1164
QY 388 GMGTLDEDPGQGMHFRFP-LRHGKYV-----SSKRMPMFMWESSASGYIR 435
Db 1165 -WTKVEKLG-----LPKIKPGOTIAFYQOYVLMFGTMARKYKWRFOAPKREYFI- 1215
QY 436 SWENNPGPQGMHSIT 450
Db 1216 NCETTPEDMSAMALS 1230

RESULT 13
RRPO_PVXHB STANDARD: PRT: 1456 AA.
ID RRPO_PVXHB STANDARD: PRT: 1456 AA.
AC O07630;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA replication protein (165 kDa protein) (ORF 1) [Contains: RNA-
directed RNA polymerase (EC 2.7.7.48); Probable helicase].
OS Potato virus X (strain HB) (PVX).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=73488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94015010; PubMed=8409947;
RA Querol M., van der Vliet R., Goldbach R., Salazar L.F.;
RT "RNA sequence of potato virus X strain HB."
RL J. Gen. Virol. 74:2251-2255(1993).
CC -1- FUNCTION: RNA-replication. The central part of this protein
CC possibly functions as an ATP-binding helicase.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X72214; CA51012.1; -
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
KW Transferase.
FT NP_BIND 735 742 ATP (POTENTIAL).
SQ SEQUENCE 1456 AA; 165138 MW; 10730A1EADA4FAB4 CRC64;

Query Match 4.4%; Score 102; DB 1; Length 1456;
Best Local Similarity 18.7%; Pred. No. 32;
Matches 104; Conservative 70; Mismatches 187; Indels 194; Gaps 26;

QY 52 EGLKGELQTRDVRGQATKILGVS-----ALGKLNSG-DATKIISG---CLDIVA 99

```


Db 255 SEILSNADADTRTKAAAG-----VELTKVLGNWNGKISQYIIAQRMAAGLSTSAAGL 309
 QY 136 VAAYIDRLSK-----HDEAIORHAAGAKRDAESSAFQVWKOOSNLTDSDLS 185
 Db 310 IASAVTLAISPLSLISADKFRANKRIEYSORFKKIDDDSLAARFKGAIDASLT 369
 QY 186 IIAANVPYKRSNFGLESRISSOGAATTSIDAKRAVDLILYCOLVYMRRETLVDLAI 245
 Db 370 TI-----STVLASYSSGIS-AAATTSLVGAPVSA-----LVGAVTGITISGLE 411
 QY 246 LYRKGAHEVASAV-----ENANRYNKEIADDTLDFLKLIPQALIGAVYHPISASETS 300
 Db 412 ASKQAMEHVASKMADVAIAEMEKKGKNGYFENGYDARHAAPLE-----DNF 457
 QY 301 KALINITYKFGVDPVPRPIGNRRKFTNSYNT 333
 Db 458 KILSOYNEKESV-----ERSVLITQOHWMT 482
 RESULT 9
 MSH5_CAEEL STANDARD; PRT; 1369 AA.
 ID MSH5_CAEEL
 AC 019272; Q9NB29;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE MSH5 protein homolog 5.
 GN MSH-5 OR F095E8.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC STRAIN-Bristol N2;
 RX MEDLINE=20469350; PubMed=11014811;
 RA Kelly K.O., Dernburg A.F., Stanfield G.M., Villeneuve A.M.;
 RT "Caenorhabditis elegans msh-5 is required for both normal and
 RT radiation-induced meiotic crossing over but not for completion of
 RT meiosis.";
 RL Genetics 156:617-630(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Percy C.M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-933 FROM N.A.
 RX MEDLINE=9900534; PubMed=9787078;
 RA Wland N.J., Panzer J.A., Kolodner R.D.;
 RT "Cloning and characterization of the human and Caenorhabditis elegans
 RT homologs of the Saccharomyces cerevisiae MSH5 gene.";
 RL Genomics 53:69-80(1998).
 CC -1- FUNCTION: Crucial component in meiotic recombination, functioning
 CC at some point after the initiation step of recombination. Plays a
 CC role in promoting the crossover outcome of meiotic recombination
 CC events. Required for formation of normal meiotic crossover, and
 CC crossover and chiasmate generated by artificially made DNA breaks.
 CC -1- SUBUNIT: HEMEROCLOMER OF MSH4 AND MSH5 (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: Germ-line specific expression.
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MSH5 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AF271389; AAF76200.1;
 DR EMBL: Z73896; CAA98059.2;
 DR EMBL: AF070070; AAC70065.1;
 DR WormPep: F09E8.3; CE26678.
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR Pfam: PF00488; Muts_C; 1.
 DR Pfam: PF01624; Muts_N; 1.
 DR ProDom: PD001263; Muts_C; 1.
 DR SMART: SM00533; Mutsd; 1.
 DR SMART: SM00534; Mutsd; 1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
 KW Meiosis; ATP-binding; DNA-binding.
 FT NP_BIND 639 646 ATP (POTENTIAL).
 SQ SEQUENCE 1369 AA; 153139 MW; DDC5FDAB7DB86C4A CRC64;
 Query Match 4.68; Score 106; DB 1; Length 1369;
 Best Local Similarity 21.78; Pred. No. 16;
 Matches 84; Conservative 44; Mismatches 165; Indels 94; Gaps 14;
 QY 35 TALKRAGSNEALEALEGL-----KGEIQTPRPGVGOATKILASV--GSALGKIN 83
 Db 777 TVLRAGGKIKYIFRMTPEGLVDCSEFALSVAKEEGIPPEVIGRACHIRALKAGTILKEIK 836
 QY 84 ---SGDATKIISGCDIYAGIATTFGPGVNGIGAVAFVSSILSLFTGSSAKNSVAAYI 140
 Db 837 AEVSNDNEKQVLEDDMDVVLADDEGF-----MAAVSEFVKRKKTSFCESSARN----V 884
 QY 141 DRALSKHDEAIAORHAAGAKRDAESSAFIOVMKOOSNLTDSDIIANVPYKRSNFI 200
 Db 885 SEEIEKERSFASSTPAKRSRSTITARSNS---VLSSRSNAAVQDLSVLDALPKKKKKVYT 941
 QY 201 G-QLESRISSOGAATTSIDAKRAVDLILYCOLVYMRRETLVDLAIYRKGAHEVASAV 259
 Db 942 GSSMESMSPPFOEDEGTGEEDOI-----SAPVSRPTL-----PSVQYASEE 987
 QY 260 ENANRYNKEIADDTLDFLKLIP--EQALIGAVYHPISASETS----- 300
 Db 988 EKQOSINRHSSTTAIHPIPIOMGEAGYKRRRSTSTSPGSAKSVTEVFKKTP 1047
 QY 301 ---KALINITYKFGVDPVPRPIGNRRKFTNSYNTYSCSEAYMGNTMFGCSNVRNP 356
 Db 1048 NVKESQVLETPPKQLSISSELEP-----KPPSEKDVLSRVSERYL----- 1087
 QY 357 NIRVSKMSDGYTMENSRKRLIYTK 383
 Db 1088 -----QSDPFRT--PISDRRSQOSSRH 1107
 RESULT 10
 FAS3_RHOFA STANDARD; PRT; 312 AA.
 ID FAS3_RHOFA
 AC P46375;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Hypothetical 33.6 kDa protein in fasciation locus precursor (ORF3).
 GN FAS3.
 OS Rhodococcus fascians.
 OG Plasmid PF1D188.
 OC Actinobacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
 CX NCBI_TaxID=1628;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D188;
 RX MEDLINE=94222824; PubMed=8169198;
 RA Crespi M., Veresche D., Temmerman W., van Montagu M., Desomer J.;
 RT "The fas operon of Rhodococcus fascians encodes new genes regulated
 RT for efficient fasciation of host plants.";
 RL J. Bacteriol. 176:2492-2501(1994).
 CC -1- FUNCTION: THE FAS-OPERON ENCODES GENES INVOLVED IN CYTOKININ

FT LIPID 563 563 PALMITATE (BY SIMILARITY).
 FT LIPID 689 689 PALMITATE (BY SIMILARITY).
 FT VARIANT 6 6 A -> T (IN STRAIN 2001).
 SQ SEQUENCE 1023 AA; 109867 MW; 196D5C0A9A2B54D CRC64;

Query Match
 Best Local Similarity 4.7%; Score 108.5; DB 1; Length 1023;
 Matches 82; Conservative 65; Mismatches 163; Indels 113; Gaps 17;

24 SDVNSL---LTKVETALKEASGSNEALELEKGEIOTKP---DRVGOATKILGSGSA 78
 199 NNWNSFQQLNKILGSLVLSNT-----KHLNKGKLNQMLPMLDNLGAGLDIVSGILSA 230
 79 LGR---LNSGDA---TKIISGCLDIAGIATTEGPGMGIGAVASVSSILSLFTGSSA 132
 251 ISASFILSNADADTGKAAAG-----VELTTKVLGNVNGKISQIYIIAQAAGLSLSA 305
 133 KNSVAVIDRALSK-----HDEAIQRAAGAKRDPASSAFIOYMKOOSNLTDG 182
 306 AGLIAVWVLAISPLSELSIADKFRANKIEEYSORFKLGYDGSLLAFHKEGTADA 365
 183 DSIILANPVYKFSNFIQLESRIISOGAATSLSDAKRAVDLILYCOLVYRETLND 242
 366 SLF-----RISTVLASVSSGIS-AAATISLVGAPVSA-----LVGAVTGIIISG 407
 243 LALYKKGNAEHAASAV-----ENANRVNKLADTLDELKLIPECALIGAVYHPISAS 297
 408 ILEASKQMEFHAASKMADIAEMWKEKHGKNYFENGCDARHAFLE----- 453
 298 ESKALANTTKYGPVDPVRPIGNKRYKFTNSWNTYSISCEAYMGNYFRCGSNVNRPN 357
 454 DNEKILISQYNKEYSV-----ERSVLITQOQMDYL----- 482
 358 IRYSKMSDEFTYEWNSDRR---KLYTKHDOGSGMGLTDEDDQGHMRFPLRHGKYW 414
 483 --IGELAG---VTRNDCKILSGSYIDYIEG---KRLKKPDEFOKQYFDPLKGNIDLS 534
 415 SSR 417
 535 DSK 537

RESULT 8
 HLYA_ECOLI STANDARD; PRT; 1024 AA.
 ID HLYA_ECOLI
 AC P08715;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemolysin, plasmid.
 GN HLYA.
 OS Escherichia coli.
 OG Plasmid Inci2 PHLY152.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hess J., Wels W., Vogel M., Goebel W.;
 RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
 its comparison with a corresponding chromosomal hemolysin sequence.";
 RL PEMS Microbiol. Lett. 34:1-11(1986).
 RN [2]
 RP PALMITOYLATION OF LYS-564 AND LYS-690.
 RA MEDLINE=95099325; PubMed=7801126;
 RA Stanley P., Packman L.C., Koronakis V., Hughes C.;
 RT "Fatty acylation of two internal lysine residues required for the
 toxic activity of Escherichia coli hemolysin.";
 RL Science 266:1992-1996(1994).
 RN [3]
 RP PALMITOYLATION OF LYS-564 AND LYS-690.
 RA MEDLINE=96404790; PubMed=8808931;
 RA Ludwig A., Garcia F., Bauer S., Jarichau T., Benz R., Hoppe J.,

RA Goebel W.;
 RT "Analysis of the in vivo activation of hemolysin (HlyA) from
 Escherichia coli.";
 RL J. Bacteriol. 178:5422-5430(1996).
 CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
 CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
 CC DEFINED.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
 CC ACTIVITY.
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN.
 CC -1- PTM: PALMITOYLATED BY HLYC. THE TOXIN ONLY BECOMES ACTIVE WHEN
 CC MODIFIED.
 CC -1- DISEASE: THE HEMOLYSIN OF E. COLI IS PRODUCED PREDOMINANTLY BY
 CC STRAINS CAUSING EXTRAINTestinal INFECTIONS, SUCH AS THOSE OF THE
 CC URINARY TRACT.
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL: M14107; AAA98233.1;
 DR InterPro: IPR001343; Hemlysn_Ca_bind.
 DR InterPro: IPR003355; RTXtoxin_N.
 DR InterPro: IPR003995; RtxA.
 DR Pfam: PF00353; hemolysinCbind; 6.
 DR Pfam: PF02382; RTX; 1.
 DR PRINTS: PR00313; CABDNCRPT.
 DR PRINTS: PR01488; RTXTOXINA.
 DR PROSITE: PS00330; HEMOLYSIN_CACIUM; 4.
 DR PROSITE: PS00330; CYTOTOXIN; 4.
 KW Hemolysin; Toxin; Cytolysin; Cytotoxin; Repeat; Calcium;
 KW Transmembrane; Lipoprotein; Plasmid.
 FT TRANSMEM 238 360
 FT TRANSMEM 268 327
 FT TRANSMEM 365 411
 FT DOMAIN 724 870
 FT REPEAT 724 729
 FT REPEAT 723 738
 FT REPEAT 742 747
 FT REPEAT 751 756
 FT REPEAT 760 765
 FT REPEAT 769 774
 FT REPEAT 778 783
 FT REPEAT 787 792
 FT REPEAT 796 801
 FT REPEAT 807 812
 FT REPEAT 817 822
 FT REPEAT 826 831
 FT REPEAT 835 840
 FT REPEAT 844 849
 FT REPEAT 856 861
 FT REPEAT 865 870
 FT LIPID 564 564
 FT LIPID 564 564
 FT LIPID 690 690
 SQ SEQUENCE 1024 AA; 110201 MW; 83944917F76C945B CRC64;

Query Match
 Best Local Similarity 4.7%; Score 108.5; DB 1; Length 1024;
 Matches 69; Conservative 52; Mismatches 139; Indels 73; Gaps 12;

24 SDVNSLTKVETALKEASGSNEALELEKGEIOTKP---DRVGOATKILGSGSALGK 81
 200 NNWNSFQQLNKILGSLVLSNT-----KHLNKGKLNQMLPMLDNLGAGLDIVSGILSA 254
 82 ---LNSGDA---TKIISGCLDIAGIATTEGPGMGIGAVASVSSILSLFTGSSAKNS 135

RESULT 5

PMPE_CHLTR	STANDARD	PRT	964 AA
ID	PMPE_CHLTR		
AC	084877		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Probable outer membrane protein pmpe precursor (Polymorphic membrane protein E).		
GN	PMPE OR CT869.		
OS	Chlamydia trachomatis.		
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=813;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-D/WM-3/CX;		
RX	MEDLINE=99000809; PubMed=9784136;		
RA	Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;		
RT	"Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis."		
RL	Science 282:754-759(1998).		
CC	-1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES) (POTENTIAL).		
CC	-1- SIMILARITY: BELONGS TO THE PMW OUTER MEMBRANE PROTEIN FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; AE001360; AAC68467.1.		
DR	InterPro; IPR003358; Chlamydia_PMP.		
DR	InterPro; IPR003357; OMP.		
DR	Pfam; PF02385; OMP; 1.		
DR	Pfam; PF02415; DUF145; 1.		
KW	Outer membrane; Signal; Multigene family; Complete proteome.		
FT	SIGNAL 1 18 POTENTIAL.		
FT	CHAIN 19 964 PROBABLE OUTER MEMBRANE PROTEIN PMPE.		
SO	SEQUENCE 964 AA: 104703 MW: 18998A7D2E571CE2 CRC64;		
Query Match	4.8%; Score 112.5; DB 1; Length 964;		
Best Local Similarity	21.3%; Pred. No. 3.6;		
Matches 110;	Conservative 61; Mismatches 184; Indels 161; Gaps 28;		
OY	21 GKRSDV---NSLTKVETALKASGSNEALELEGKELQTPDRVGATKILGSVG 76		
DB	254 GNRGINVFNNRCFKNVERASEAS-----DGAIKVTTRLDVTGNGRIFES-- 301		
OY	77 SALGLNSGDATKRTISGLDIYAGIATPGCGVYGICGAVASVSIISLFTGSSAKKSV 136		
DB	302 -----DNTTKYGGAI---YAPVYLVVNGPYFINNNANKG 336		
OY	137 AAVIDRALSKHDEAIORHA-----AGAKR---DEFASSAFIOY 172		
DB	337 GAIYIDGTSNSKISA-DHRAIFENNIYNTNANGTSTSNAPPRNMTAVASSGEILL 395		
OY	173 MKQOS-NLTSD-LSIIAANPVYKSNFTGLESRIISOGAATISLSPAKRAVD----- 224		
DB	396 GAGSQNLIFFYDPIEVSNAGSV-SFNKEADQTSVFSGATVNSADEFQRMLQTKTPAP 454		
OY	225 -----FILL--YCOLVYMRETL--LVDL--AIL--YRKGNAHVASAVENARVYKE 268		
DB	455 LTLNSGFLICIEHMOIYNNRFOTGCVSLGNGATLSTCKNKTGD-----SASNAITLKH 510		
OY	269 LAADTLDFLHLKIPQALIGAVYHPIISASETSKAILANT---KRYGVDPVPRP---G 320		
DB	511 IGLN-----LSSITLKGAEI-----PLWVEPTNNSNNYTAADAAVFSLSDVLTSLIDYG 561		

RESULT 6

ACA9_ARATH	STANDARD	PRT	1073 AA
ID	ACA9_ARATH		
AC	091041		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Potential calcium-transporting ATPase 9, plasma membrane-type		
DE	(EC 3.6.3.8) (Ca2+-ATPase, Isoform 9).		
GN	ACA9 OR AT3G21180 OR MXL8.3.		
OS	Arabidopsis thaliana (mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Columbia;		
RX	MEDLINE=20277480; PubMed=10819329;		
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;		
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and TAC clones."		
RL	DNA Res. 7:131-135(2000).		
CC	-1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE TRANSLLOCATION OF CALCIUM FROM THE CYTOSOL OUT OF THE CELL OR INTO ORGANELLES (BY SIMILARITY).		
CC	-1- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+)(in) = ADP + phosphate + Ca(2+)(out).		
CC	-1- ENZYME REGULATION: ACTIVATED BY CALMODULIN (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- DOMAIN: THE N-TERMINUS CONTAINS AN AUTOINHIBITORY CALMODULIN-BINDING DOMAIN, WHICH BINDS CALMODULIN IN A CALCIUM-DEPENDENT FASHION (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2 ATPASES). SUBFAMILY 1IB.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; AB023045; BAB01709.1. -		
DR	HSSP; P04191.1EUL.		
DR	InterPro; IPR001757; ATPase_E1-E2.		
DR	InterPro; IPR004014; Cation_ATPase.		
DR	InterPro; IPR001454; Hlgase/hydrlase.		
DR	Pfam; PF00122; E1-E2_ATPase; 1.		
DR	Pfam; PF00689; Cation_ATPase_C; 1.		
DR	Pfam; PF00690; Cation_ATPase_N; 1.		
DR	Pfam; PF00702; Hydrolase; 1.		
DR	PRINTS; PR00119; CATATPASE.		
DR	TIGRFAMS; TIGR01116; Ca_ATPase; 1.		
DR	PROSITE; PS00154; ATPASE_E1_E2; 1.		
KW	Hydrolase; Calcium transport; Transmembrane; Phosphorylation; ATP-binding; Metal-binding; Magnesium; Calmodulin-binding;		

Db 61 KPRVQATKILGVSALGKLSGDAITIISSCLDIIVAGIATTEGCGVIGAVASFV 120
 QY 121 SSILSFTGSSAKNSVAVIDRALSKHDEAIORHAGAKRDEFAESSAFIOVMQOENLT 180
 Db 121 SSILSFTGSSAKNSVAVIDRALSKHDEAIORHAGAKRDEFAESSAFIOVMQOENLT 180
 QY 181 DSDLSIAANVPYKFSNFQGLSRSISOGAATTSLSDAKRAVDIFILLYCOLVYMRTEL 240
 Db 181 DSDLSIAANVPYKFSNFQGLSRSISOGAATTSLSDAKRAVDIFILLYCOLVYMRTEL 240
 QY 241 VDIALIRKGNAEHVAHAVENANVNKELADTLDLFLHKLIPDALIGAYHPIASETS 300
 Db 241 VDIALIRKGNAEHVAHAVENANVNKELADTLDLFLHKLIPDALIGAYHPIASETS 300
 QY 301 KALINTKRYGVDVPRPIGNRRYKFTNSYMTYISCSAYMGNYMFGCSNVNPNIRV 360
 Db 301 KALINTKRYGVDVPRPIGNRRYKFTNSYMTYISCSAYMGNYMFGCSNVNPNIRV 360
 QY 361 SKMSDGYTWMNSDRKLYTKHDQGWGCTLDEPDQGHMRFILRHGKYMVSRRMP 420
 Db 361 SKMSDGYTWMNSDRKLYTKHDQGWGCTLDEPDQGHMRFILRHGKYMVSRRMP 420
 QY 421 NMFMYESSASGYIRSMENPNPGOHWSIT 450
 Db 421 NMFMYESSASGYIRSMENPNPGOHWSIT 450

RESULT 2

CTXA_CARAL STANDARD: PRT: 463 AA.

AC Q9GN8; 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Toxin-A precursor (CatX-A).
 GN CAT-1.
 OS Carybdea alata (Hawaiian box jellyfish).
 OC Eukaryota; Metazoa; Cnidaria; Cubozoa; Cubomedusae; Carybdeidae;
 OC Carybdea.
 OX NCBI_TaxID-112899;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 158-168 AND 431-448.
 RC TISSUE=Cnidoblast;
 RX MEDLINE-20422302; PubMed-10964708;
 RA Nagai H., Takawa K., Nakao M., Sakamoto B., Crow G.L., Nakajima T.;
 RT "Isolation and characterization of a novel protein toxin from the
 Hawaiian box jellyfish (sea wasp) Carybdea alata.";
 RL Biochem. Biophys. Res. Commun. 275:589-594(2000).
 CC -1- FUNCTION: Has potent hemolytic activity. lethally toxic to
 crayfish when administered via intraperitoneal injection (LD50 =
 5-25 microg/kg). Causes cutaneous inflammation in humans.
 CC -1- PPM: There are disulfide bonds.
 CC -1- SIMILARITY: BELONGS TO THE JELLYFISH TOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB036714; BAB12727.1;
 KW Toxin; Hemolysis; Signal.
 FT SIGNAL 1 18
 CHAIN 19 463
 TOXIN-A
 SEQUENCE 463 AA: 51604 MW: 461253DFCEB9D3C0 CRC64;

Query Match 42.7%; Score 992.5; DB 1; Length 463;
 Best Local Similarity 43.5%; Pred. No. 2.2e-61;
 Matches 197; Conservative 95; Mismatches 148; Indels 13; Gaps 7;

QY 9 LFTVALTSARKG-----KRSVNSLITKVELALKENS--GSNEAL-EALEGLKEIQ 60
 Db 11 LFTVALTSARKG-----KRSVNSLITKVELALKENS--GSNEAL-EALEGLKEIQ 60
 QY 61 KPRVQATKILGVSALGKLSGDAITIISSCLDIIVAGIATTEGCGVIGAVASFV 119
 Db 71 EPATTTAKVSTIVGSGVSLSKFSGDPPFVAGSCIDIIVSVATTFGPGYGAIVASL 130
 QY 120 VSSILSFTGSSAKNSVAVIDRALSKHDEAIORHAGAKRDEFAESSAFIOVMQOENLT 179
 Db 131 SSILSFTGSSAKNSVAVIDRALSKHDEAIORHAGAKRDEFAESSAFIOVMQOENLT 190
 QY 180 TDSLSIAANVPYKFSNFQGLSRSISOGAATTSLSDAKRAVDIFILLYCOLVYMRTEL 239
 Db 191 TEVHDSVDAVAVDAFTNMLGVLSEIRINRGSVTDNNEMKRTNIFLILQSLVRETL 250
 QY 240 VDIALIRK--GNAEHVAHAVENANVNKELADTLDLFLHKLIPDALIGAYHPIAS 297
 Db 251 LTVQVILLYRAGAYDELALSLTSDQNKKEATRETIVTFLHQMETKYSICGSYYPIIDS 310
 QY 298 ETSKALINTKRYGVDVPRPIGNRRYKFTNSYMTYISCSAYMGNYMFGCSNVNPNIRV 356
 Db 311 KAIGILKLTFFGVDVPRPIGNRRYKFTNSYMTYISCSAYMGNYMFGCSNVNPNIRV 370
 QY 357 NIKVSKSDGYTWMNSDRKLYTKHDQGWGCTLDEPDQGHMRFILRHGKYMVSRRMP 416
 Db 371 GIRIKLENGYHTI-ILSKAMVTYKHAOGWGCTLDEPDQGHMRFILRHGKYMVSRRMP 429
 QY 417 KRPVQATKILGVSALGKLSGDAITIISSCLDIIVAGIATTEGCGVIGAVASFV 449
 Db 430 KRPVQATKILGVSALGKLSGDAITIISSCLDIIVAGIATTEGCGVIGAVASFV 462

RESULT 3

CTXA_CHIOU STANDARD: PRT: 462 AA.

AC P58762; 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Toxin-A precursor (CqTx-A).
 GN CQT-A.
 OS Chirospalmus quadrifidus (Box jellyfish).
 OC Eukaryota; Metazoa; Cnidaria; Cubozoa; Cubomedusae; Chirodroplidae;
 OC Chirospalmus.
 OX NCBI_TaxID-130731;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cnidoblast;
 RX MEDLINE-21854932; PubMed-11866126;
 RA Nagai H., Takawa K., Nakao M., Oshiro N., Iwanaga S.,
 RT "A novel protein toxin from the deadly box jellyfish (Sea Wasp,
 Habu-kurage) Chirospalmus quadrifidus.";
 RL Biosci. Biotechnol. Biochem. 66:97-102(2002).
 CC -1- FUNCTION: Has potent hemolytic activity. lethally toxic to
 CC crayfish when administered via intraperitoneal injection (LD50 =
 80 microg/kg). Causes cutaneous inflammation in humans.
 CC -1- SUBCELLULAR LOCATION: Cnidocyst and then secreted.
 CC -1- PPM: There are disulfide bonds (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE JELLYFISH TOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB045319; BAB82520.1;
 KW Toxin; Hemolysis; Signal.
 FT SIGNAL 1 19
 CHAIN 19 462
 TOXIN-A
 SEQUENCE 462 AA: 51604 MW: 461253DFCEB9D3C0 CRC64;

GenCore version 5.1.4-P5_A578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 11:14:48 ; Search time 25 Seconds

(without alignments)
746.574 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322
Sequence: 1 MILKHLPMFLFYLAITSKHK.....SGYRSMENNPGGCHMSIT 450

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swisaprot_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2322	100.0	450	1	CTRX1_CARRA
2	992.5	42.7	463	1	O9GVN72 carybdea ra
3	423.5	18.2	462	1	CTRX_CARAL
4	116.5	5.0	376	1	MDRC_YEAST
5	112.5	4.8	964	1	PMPE_CHLTR
6	110	4.7	1073	1	ACN9_ARATH
7	108.5	4.7	1023	1	HUY1_ECOLI
8	108.5	4.6	1024	1	HUYA_ECOLI
9	106	4.5	312	1	FAS3_RHOA
10	105.5	4.5	506	1	ABR_ECOLI
11	105	4.5	506	1	RRPO_PVACP
12	103	4.4	1456	1	RRPO_PVACP
13	102	4.4	1456	1	RRPO_PVACP
14	101.5	4.4	299	1	PMAS_ARATH
15	101	4.3	393	1	FLAA_PSEAE
16	101	4.3	535	1	HTRL_HAANI
17	101	4.3	1520	1	ACPD_VIRCH
18	100.5	4.3	909	1	SBCC_DEIRA
19	100	4.3	437	1	YABE_BACSU
20	100	4.3	444	1	Y808_CHLPP
21	99	4.3	948	1	PMAS_ARATH
22	99	4.3	1935	1	MYSS_CYPCA
23	98.5	4.2	379	1	FLAA_VIRCH
24	97.5	4.2	487	1	SR54_ENTHI
25	97	4.2	825	1	GNJ3_BACSA
26	97	4.2	1934	1	MYH7_MESAV
27	97	4.2	1935	1	MYH7_RAT
28	96.5	4.2	1065	1	NOHG_RHIME
29	95.5	4.1	3027	1	POG6_PXPV1
30	95	4.1	381	1	YBD2_YEAST
31	95	4.1	434	1	T147_HUMAN
32	95	4.1	483	1	KICL_MOUSE
33	95	4.1	675	1	ATKB_DEIRA

34	95	4.1	1120	1	STFR_ECOLI	P76072 escherichia
35	95	4.1	1456	1	RRPO_PVX	P09395 potato viru
36	94.5	4.1	454	1	SR54_AQUAE	O67615 equifex aeo
37	94.5	4.1	548	1	HLVB_VIRCH	P15492 vibrio chol
38	94.5	4.1	1706	1	CYVA_BORPE	P15318 bartonella
39	94	4.0	374	1	FLAI_BABA	P35633 bartonella
40	94	4.0	450	1	ADFP_BOVIN	O9TUM6 bos taurus
41	94	4.0	557	1	MCPS_ENTAE	P21822 enterobacte
42	94	4.0	1935	1	MYH7_HUMAN	P12883 homo sapien
43	93.5	4.0	427	1	HISX_HAETI	P44001 haemophilus
44	93.5	4.0	580	1	IPAB_SHIFT	P18011 shigella fi
45	93.5	4.0	903	1	MSPI_SCHPO	P87320 schizosacch

ALIGNMENTS

RESULT 1	ID	CTRX1_CARRA	STANDARD:	PRT:	450 AA.
AC	O9GVN72	CTRX1_CARRA			
DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Toxin-1 precursor (CTRX-A / CTRX-B).				
GN	CRT-1				
OS	Carybdea rastoni (Box jellyfish).				
OC	Eukaryota; Metazoa; Cnidaria; Cubozoa; Cubomedusae; Carybdeidae;				
OC	Carybdea.				
OX	NCBI_TaxID=78582;				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Cnidoblast;				
RX	MEDLINE=20422301; PubMed=10964707;				
RA	Nagai H., Takawa K., Nakao M., Ito E., Miyake M., Noda M., Nakajima T.;				
RT	"Novel proteolaceous toxins from the box jellyfish (sea wasp) Carybdea rastoni".				
RL	Biochem. Biophys. Res. Commun. 275:582-588(2000).				
CC	-1- FUNCTION: Has potent hemolytic activity. Lethally toxic to mice (LD50 = 20 microg/kg via i.v.) and grayfish (LD50 = 5 microg/kg via intraperitoneal injection). Causes cutaneous inflammation in humans.				
CC	-1- SUBCELLULAR LOCATION: Cnidocyst and then secreted.				
CC	-1- TISSUE SPECIFICITY: Synthesized in the tentacle then migrates to the cnidocyst.				
CC	-1- PTM: There are disulfide bonds.				
CC	-1- SIMILARITY: BELONGS TO THE JELLYFISH TOXIN FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
DR	EMBL: AB015878; BAB12728.1;				
KW	Toxin; Hemolysis; Signal.				
FT	SIGNAL 1 18 POTENTIAL.				
FT	CHAIN 19 450 TOXIN-1.				
SO	SEQUENCE 450 AA; 49392 MW; CD393CF25BFD2FD CRC64;				
QY	Query Match	100.0%	Score 2322;	DB 1;	Length 450;
QY	Best Local Similarity	100.0%	Pred. No. 2.8e-153;		
QY	Matches 450; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
DB	1 MILKHLPMFLFYLAITSKHKGRSDVNSLTVERALKEASGNSPALEALGELGIOT 60				
QY	1 MILKHLPMFLFYLAITSKHKGRSDVNSLTVERALKEASGNSPALEALGELGIOT 60				
QY	61 KEDRYGOATKILGVSALGKLSNGDATKIIISGCDIYAGIATPFGPGVMGIGAVASPV 120				

QY 89 KIISGCDIYAGIATTEGPGVGMIGAV-----ASPVSSILSLFTG--- 129
 Db 2157 KSESDYQDANADQNNANDAVTNAEGITISATNNPEMNPDTINCKASVNSKSLNDEK 2216
 QY 130 -SSAKNSVAVIDRALSKHRDEAIQRHAAGAKRDEAFESSAFIQVMKOS----- 177
 Db 2217 LAAKQTAKSDIGRLTDLNN---AQRTANAEDQANPLAAVTAKNKATSLNTAMGNLK 2273
 QY 178 -----NLDS-----LSITAAVPPYKFSNFTQLESRIISOGAATTSLSDAK 220
 Db 2274 HALAEKDNTRKRSVNTDADQPKQOAYDTAVTQAEAITNANGSNANETQVOAALNQLNAK 2333
 QY 221 RAVDFTLLYCOLVYMRFTLLVDLAILYRKGNAEHVA--SAVENANRYNKELEA-DTLDL 277
 Db 2334 NDINGDNKVAO---AKETARALASYSNLNNAOSTAATSOIDNATYADVTAAONTANEL 2390
 QY 278 HKLIPEQALIGAVYHPISASETSKALINTYKFGVDPVPRPIGNRRYKFTNSYNTYSIC 337
 Db 2391 N-----TANGOLONGINDQNTYKQOVNFT-----DADQ---GKRDATYNAVTNAQGITL 2435
 QY 338 SEA 340
 Db 2436 DKA 2438

RESULT 15

A72428
 methyl-accepting chemotaxis protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: A72428
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MID:99287316; PMID:10360571
 A:Accession: A72428
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-656 <ARN>
 A:Cross-references: GB:AE001690; GB:AE000512; NID:g4980496; PIDN:AAD35117.1; PID:g498050
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0023
 C:Superfamily: probable methyl-accepting chemotaxis transducer

Query Match 4.8%; Score 111.5; DB 2; Length 656;
 Best Local Similarity 20.4%; Pred. No. 5.7; Mismatches 107; Indels 95; Gaps 10;
 Matches 63; Conservative 44;

QY 24 SDVNSLLTKVETALKEASGSENALEALGL--KGEIOTRPPDRYGQATKILGSVGSALGK 81
 Db 393 NNMSSALTEVTSGEVEFAAASQNIKTQDLTERSEAVTKAARG--TERVEAVGVYINK 450
 QY 82 LNSG-----DATIISGCDIVAGIA-----TTFGPGVGMIGAV 116
 Db 451 LKGSARORDYLRELYDSAKTIGETIVDTISSIAEOTNLALNAIEAARAGEAGRGFAVV 510
 QY 117 ASFVSSILSLFTGSSAKNSVAVIDRALSKHRDEAIQRHAAGAKRDEAFESSAFIQVMKQ 176
 Db 511 ADEIRKI-----AESQRAETEDIAKKMLSSLRATTEHVENGSKEMFEGVDEIAYNGEE 562
 QY 177 SNLTDSDLSITAAVPPYKFSNFTQLESRIISOGAATTSLSDAKRAVDFTLLYCOLV 233
 Db 563 -----VTKRREITGRIEETINSMTENTAAATAOEAGAA----- 594
 QY 234 VMRETLVDLAILYRKGNAEHVA SAVENANRYNKELEA DTLDL FHLKLIPEQALIGAVYHP 293
 Db 595 -----AEMASAMDNVTKI-VEGVESLNRMSLEIDQT----- 627
 QY 294 ISASETSKA 302
 ||: ||:

Db 628 ESARVSEA 636

Search completed: May 19, 2003, 11:17:21
 Job time: 28 secs

Query Match 4.88; Score 112.5; DB 2; Length 964;

[illegible]

A:Gene: TM1143
C:Superfamily: methyl-accepting chemotaxis protein
Query Match 5.0%; Score 116; DB 2; Length 530;
Best Local Similarity 19.4%; Pred. No. 2.1;
Matches 54; Conservative 63; Mismatches 122; Indels 40; Gaps 9;
QY 15 ITSAKHGRSDVNSLTFTVET---ALKEASGSNEALEBGLKEIQTKPRVQAKRI 71
DB 248 ITNOLGISTEKMDNSTRIEISASVOETTAGSEISATKNIADSAQOASPADOSTOL 307
QY 72 LGSVSAIGKINSGDATKISCLDIYAGIATTFGGPVMGIGAVSFYSILSTFTGSS 131
DB 308 AKEDADALKTY--EYTRKISNSAKDVERVVSF---OKGBEITTSFETINAI----- 356
QY 132 AKNSVAAYIDRALSKHRDEAIOHRAAGAKRDPFAESSAPIQVKQOOSNLTDSDLSITANV 191
DB 357 AEQTLLALNALIEAR-----AGEAGRGFAVVADEIRKLAESQASENVRVYNEI 409
QY 192 ----PYKFSNFICQLESRIQGAATTSLSDAK-----RAVDFTLLYCOLVYMRET 238
DB 410 RSIADAGKVS---EITARVEEGTKLADDEKINSIVGAVERINEMQNTAAALEQT 466
QY 239 LLYD---LAILYRKGNAEHVASAVENANVRKEIADT 273
DB 467 AAVDEITTRMTENAKNAEITNSVKEVNAKLOEISAST 504

RESULT 10

AC1647
Lactobacillus phage phi-gle minor capsid protein 1608 homolog 11n1716 [Imported] - Liste
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC1647
R:Glasner, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlan, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Meok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1647
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1571 <GLA>
A:Cross-references: GB:AL59022; PIDN:CAC96947.1; PID:g16414203; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: 11n1716

Query Match 5.0%; Score 116; DB 2; Length 1571;
Best Local Similarity 20.3%; Pred. No. 8.9;
Matches 84; Conservative 74; Mismatches 148; Indels 108; Gaps 18;
QY 11 IVLAITSANHGKRSVDNSLTFTVETALKEASGSNEALEBGLKEIQTKPRVG--QA 68
DB 141 LVVAGKAFENYASDPTNESINKVDVAFKNA-----ESVKKWSKTTLLNIGLAQG 190
QY 69 TKI-----LGSVSAIGKINSGDATKISCLDIYAGIATTFGGPVMGIGAVASFV--- 120
DB 191 TALDLAATYGDMSTSMG--ISTOEAKMSTSMVLA-----GNLASRKNID 234
QY 121 ----SSILSLFTG--SSAKNSVAVIDR-----ALSKHRDEAIOHRAAGAKRDPFAESSA 168
DB 235 IDRAANALNGVFTGETEALSLGIVMTQTLNQFALETGAGKVAKSSTEYTKONIAE-- 292
QY 169 FIVVMQOOSNLTDSDLSITANVPYKFSNFICQLESRIQGAATTSLSAKKAVDFILL 228
DB 293 ----KAOKRLNBAIKREHGKNSLEAPQOKLOEIOAKTSEG--AKVNLKMKO--DELVR 344
QY 229 YCOLVYMRFTLVDLAILYRKGNAEHVASAVENANVR---NRELADTLDLFLHKLIPQ 284

DB 345 LRYNVYMKOT-----TNAGDEFKNTSDQANASRVFSESSEKELASNAGQFLPLI--T 395
QY 285 ALICAVYPIASSENKAILNTTKYFGVDPVPRIGNRRYKFTNSYMWTFYSGSPAYM-- 342
DB 396 PLI-----IQADEFKTL-----SDIPGAVKGMKEKFKP--EVFEVGDFFPKD 439
QY 343 -----GNMYFRCNSVNRPNIRSKMSDGFYTMENSDRK 377
DB 440 LIPSAKELAKSMGPGFIEGVLAFKGLIVLNTVTYIPAFKAVTKTRRNPDSMK 493

RESULT 11

AG0438
Probable Rns accessory genetic element YP03606 [Imported] - Yersinia pestis (strain C
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AG0438
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
ll, M.; Rutherford, K.; Simmonds, N.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrer
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0438
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-800 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC93075.1; PID:g15981527; GSPDB:GN00175
C:Genetics:
A:Gene: YP03606

Query Match 5.0%; Score 115; DB 2; Length 800;
Best Local Similarity 17.3%; Pred. No. 4.2;
Matches 101; Conservative 79; Mismatches 202; Indels 202; Gaps 24;

QY 12 VLAITSAKHGRSDVNSLTFTVETALKEA---SGSNEAL-----EALEGKEIQTRP 62
DB 1 MFAHDKANNKNAADKGTGVTPTIHVSPEAPYVSGAQMTEAGIGDVGATGAGK----- 55
QY 63 DRVQATKTLGVSGLKGLNSGDAT-----KIISGC--LDIVAGIATTFGGPVG 110
DB 56 -RINDAAEYAKTALDAKAVLDGGVTPNIVATGPAPTYVSGAQMTEAGIGDVGATG 114
QY 111 MGIGAVASFVSSILSLFTSSAKNSVAAYIDRALSKHRDEAIOHRAAGAKRDPFAESSAP 170
DB 115 KRINDAAE-----YAKTLEAKAKVLDGGVTPM--LAAGGAGGLPDSAAATIS 162
QY 171 QVAKQOOSNLTDSDLSITANVPYKFSNFICQLESRIQGAATTSLSAKRAVDFTLL 229
DB 163 RLTKOPSGL---QFTLTASLPQTFPAVDFLTSEKLSPPVLNVLGASADPAVDPAVL 219
QY 230 ---COLVYMR-----TLVLDLAILYRK 250
DB 220 DEDATLFIWRGEVLAISITGVNASFEQDGTGFHQRYSWVIRPALWRTSLRRNARIFQOA 279
QY 251 NAEHVASAVENANVRVK-----ELAADLDFLHKLIPQAL----- 286
DB 280 SVEEITTLKEKINDFAFGFRHPHPVREFCVOIESPDTLQRLTAEGIFYYFEESA 339
QY 287 -----IGAV-----YHP--ISASETSKAILNTTKYGVDPVPRPIGNRRYKFTN 328
DB 340 GKRTVYVADVGLPKGASLPYRPNVAQAQOELSITTFPRSAQVR--PAMVOLKDYTFKN 397
QY 329 STW-NTYSCS-----EAYMGNYFRCNSVNRPNIRSKMSDGF 367
DB 398 PNMVAAFSEQSGELQNRDPYEHFDPGRKKAQHODFTFRRLDLRNDANLGQASND 457
QY 368 YTMENSDRRLYTTKIDOG--WGW-----GTLDEDPGCGG-----HMFETPL 407
DB 458 FTLOPQOLSLY--NIPRODLNHAQWLLGTOHSGKQMLQASGQGTGLFHFHFEFIP- 514
QY 408 RHGKYVSSKRWPMNMTWYESSASGYIRSWENNP-----GPO 444

Matches 100; Conservative 54; Mismatches 127; Indels 174; Gaps 23;

QY 4 KHL-----PWLFIYLAITSAGHGRSDVNSLITKVFETALKEASG-SNEALEALEGLK---- 55
 DB 225 KHLQVAPM-----DVVQSGQPSISILPSEPLSSDLEKSTEAALVFCRSLKRVV 286
 QY 56 ---GEIQKPDVNGATKILGVSGLAIGLNSGDATKIISGLDIYAGIATTFGGVPGNG 112
 DB 287 FRSQLEPIPD-----SLDSEFISERLD-----STHHGAVGAG 321
 QY 113 -----IGAVASVSSILSLFTGSSAKNSVAAYDR----- 142
 DB 322 BCESTKESHSLIKKKKKKKKKKSVLATSSDASVYDSKAKN---ALVDSSESGAVREED 378
 QY 143 -----ALSKHDEAIOR-----HAAGAKRDFESSAFIOVMQKQNLNDSLS 185
 DB 379 VDHRAEAEAOACSTEKHR-EMQORLEPHEESNSSESNAARHISDRSDSDVD 437
 QY 186 IIAANVPYKFSNFTGLESRISSQAATTSLSDAKRAVDLILYQOLVYMRFTLLVDLAI 245
 DB 438 LGSA---VROLEFIPDIOER---AATV-----IRRYRDLG- 469
 QY 246 LYRKNMAEHA-----SAVENANRVNKEADTLDLHLKILPEQALIGAVYHPISASST 299
 DB 470 LFEKKAQGVAILREKFSKAE-----NKOIEKNVODFL-----SLTGIES 509
 QY 300 SKAILNTYKFGVDPVPRPIGN--RRYKFTNSYVNTYISCSBAVMGNVYRFGCSVVRPNP 357
 DB 510 ADKLITYDRY---PEKTLITLTKRKHAH-----RLHIGKGLARWKVLY---Y 551
 QY 358 IRVSMMSD-----GFTYHNSDRKRLYITKHDGM 387
 DB 552 YRAKIFDVNNYKGRYNEEDTKKLKAYHSLHGNMW 586

RESULT 8
 DEBYWC
 malate dehydrogenase (EC 1.1.1.37), cytosolic - yeast (Saccharomyces cerevisiae)
 N:Alternate names: Protein AOE423; protein 00537; protein YOL126c
 C:Species: Saccharomyces cerevisiae
 C>Date: 30-Jun-1993 #sequence_revision 19-Jul-1996 #text_change 16-Jun-2000
 A:Accession: S63444; S12937; A34986; S05770; S66823; S71982
 R:Casamayor, A.; Khalid, H.; Balcells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J.
 Submitted to the EMBL Data Library, November 1995
 A:Description: Sequencing of a 13.4 kbp fragment of the left arm of chromosome XV reveal
 A:Reference number: S63440
 A:Molecule type: DNA
 A:Accession: S63444
 A:Residues: 1-423 <CAS>
 A:Cross-references: EMBL:U41293; NID:g1209710; PIDN:ACC49466.1; PID:g1209715
 A:Experimental source: strain FY1679
 R:Minard, K.I.; McAlister-Henn, L.
 Mol. Cell. Biol. 11, 370-380, 1991
 A:Title: Isolation, nucleotide sequence analysis, and disruption of the MDH2 gene from S.
 A:Reference number: S12937; MUID:91094852; PMID:1986221
 A:Accession: S12937
 A:Molecule type: DNA
 A:Residues: 47-423 <MINI>
 A:Cross-references: EMBL:M62808; NID:g171915; PIDN:AAA34766.1; PID:g171916
 A:Accession: A34986
 A:Molecule type: protein
 A:Residues: 48-73 <MIN2>
 R:Kopetzki, E.; Entian, K.D.; Lotspeich, F.; Necke, D.
 Biochim. Biophys. Acta 912, 398-403, 1987
 A:Title: Purification procedure and N-terminal amino acid sequence of yeast malate dehyd
 A:Reference number: S05770; MUID:87185517; PMID:352052
 A:Accession: S05770
 A:Molecule type: protein
 A:Residues: 48-73, 'P', 75-81 <KOP>
 R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas, C.
 Submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66814
 A:Accession: S66823

A:Molecule type: DNA
 A:Residues: 1-423 <ARI>
 A:Cross-references: EMBL:274868; NID:g1420008; PIDN:CA99145.1; PID:g1420009; GSPDB
 A:Experimental source: strain S288C
 R:Casamayor, A.; Khalid, H.; Balcells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino,
 Yeast 12, 1013-1020, 1996
 A:Title: Sequence analysis of a 13.4 kbp fragment from the left arm of chromosome X
 ending frames.
 A:Reference number: S71978; MUID:97051588; PMID:8896265
 A:Accession: S71982
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-423 <CAW>
 A:Cross-references: EMBL:U41293; NID:g1209710; PIDN:ACC49466.1; PID:g1209715
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 199
 C:Genetics:
 A:Gene: SGD:MDH2; MIPS:YOL126c
 A:Cross-references: SGD:S0005486; MIPS:YOL126c
 A:Map position: 15L
 C:Superfamily: L-lactate dehydrogenase
 C:Keywords: cytosol; homodimer; NAD; oxidoreductase
 F:48-423/Product: malate dehydrogenase #status experimental <MAT>
 F:228,201/Active site: Asp, His #status predicted
 F:231/Binding site: substrate (Arg) #status predicted

Query Match 5.0%; Score 116.5; DB 1; Length 423;
 Best Local Similarity 20.2%; Pred. No. 1.4;
 Matches 72; Conservative 65; Mismatches 125; Indels 95; Gaps 17;

QY 13 LAITSAGHGRSDVNSLITKVFETALKEASG-----NEALEALEGLGEIOTRPD 63
 DB 62 IAILGAAGIGQSLILKLAQLOLYOLKESNSVTHILHALVDVNGEALNGVTADLSHDT 121
 QY 64 RVGQATKILGVSGLKGLNSGDATKISG-----CLDIYAGIATTFGGVPGNG 112
 DB 122 PIVSSH--SPAGGIENCLHNASIVYIPAGVRRKGMTRDDLFVNAGIISQLDSI--- 176
 QY 113 IGAVASVSSILSLFTGSSAKNSVAAYDRALSKH---RDEAIORHAAG-AKRPFAESSA 168
 DB 177 --AECCDLSKVFVLYISNPVMSVLPVWMSNLKHPGSRNGIERIRMGVTKLDIVRAST 234
 QY 169 FIYVMKQOQSNLTD-----SDLSITAN-----VPYKFSNFTGOL----- 203
 DB 235 FLREINIESGLTPRNSMPDVPVIGSHGETIIPFSQSNFLRNLNEDQLYLHRYQYG 294
 QY 204 -----ESRISGGAFTSLSDA--KRAVDLILYQOLVYMRFTLLVDLAILRKNAEHA 256
 DB 295 GDEVYKANKGKSAATLSMAHGYKCVQFV-----SLL-----GNIRQIH 335
 QY 257 SA---VENANRVNKEADTLDLHLKILPEQALIGAVYHPISASSTSKAILNTYKY 309
 DB 336 GTTYVPLKDNANPFIARGAD-----QLLP--LVDGADYFAIPLITTKGV-STVDY 383

RESULT 9
 C72291
 methyl-accepting chemotaxis protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 A:Accession: C72291
 R:Nelson, K.E.; Clayton, R.A.; Gail, S.R.; Gwin, M.L.; Dodson, R.J.; Haf, D.H.; H
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: C72291
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-530 <ARN>
 A:Cross-references: GB:AE001771; GB:AE000512; NID:g4981678; PIDN:AAD36219.1; PID:g49
 A:Experimental source: strain MSB8
 C:Genetics:

Matches 107; Conservative 55; Mismatches 173; Indels 130; Gaps 23;

QY 21 GKRSDV-NSLITKVTALKEASGSNEALELEGKEIOTKPRVGOATKIL--GSYGS 77
 Db 310 GVRDYNRSKVPKPDGKKEK--TKEDADNPKDALGANLGDITSTSHAAMLAKAGVGS 367
 QY 78 ALGLNLSGDAKTIISGLDVIATFTFGVGVGIGAVASVSIISLTFGSSAKNSVA 137
 Db 368 ALGKATG-----LAGMAA--GAGLGPAMVAGSOGSALGAPALAGNSVA 413
 QY 138 AVIDR--ALSKHDEAIORHAAGAKRDESSAFIOVKQOONLTDLSL--IIAANPV 193
 Db 414 AVGGAAVALGKHIGSIKKGADTVSNLNGES-----YPELTDIEDIADLATKDPEN 464
 QY 194 YKSNFTIGLESRISSGAAATSLSD-----AKRAVDLILYC-----QVYVRETLVDL 243
 Db 465 WKANPNSVAASRLKQAFPNASDAEIAKKVAKTNSDQMSRTNORODLQNMKKT---- 519
 QY 244 AILYKGNNAEIVASAVENANRNKELADTLDFHKLIPQALIGAVYHPIASPTSK-- 301
 Db 520 ATPY--GNARLDVNAATNA--FOKGYEGDHKDTFMSQLPEN-----MSAEKKEKQW 566
 QY 302 -----AILNY-----TKYFGVP--DVPRPIGNRRYKFTNSYWTYSICSEAYMGNYM 346
 Db 567 NDHLNAKVOGFRNHAEDQATKAGAMPYDAKDKQNN--LFDKSYVNKDAFASOLALG--- 621
 QY 347 FRGCSNVRNPNIRYSKSDGF--YTMENSDR-----RKLYITHKDGWG 388
 Db 622 -----KVGKAVEGVKGETLESQHONMGVAGLWASTAFAFKGTADHAKGF- 668
 QY 389 WGTLEDDEPGDGHMRFIPLRHGYWNSKRWPMFMWMESSASGY 433
 Db 669 ---MKQFPAD-----MSQOEKELANKKHLDQKVOGF 696

RESULT 5
 AE0150
 Probable methyl-accepting chemotaxis protein tsr2 [imported] - Yersinia pestis (strain C)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C:Accession: AE0150
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tildall, R.W.; Holden, M.T.G.; Prentice, M.B.; demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11565360
 A:Status: preliminary
 A:Accession: AE0150
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-595 <KUR>
 A:Cross-references: GB:AL509842; PIDN:CAC90064.1; PID:g15979284; GSPDB:GN00175
 C:Genetics: tsr2
 A:Gene: tsr2

Query Match 5.1%; Score 119.5; DB 2; Length 595;
 Best Local Similarity 23.8%; Pred. No. 1.4; Indels 41; Gaps 11;
 Matches 57; Conservative 39; Mismatches 103;

QY 21 GKRSDVNSLITKVTALKEASG-----SNEALELEGK--GEIOTKPRDRVGO 67
 Db 371 GRTSAAASLQOTSALEQISATVAOSASARANNVAFSSSDASRGDVITK----- 424
 QY 68 ATKILGVSAGALKLNSGDAKTIISG-----CIDIYAGIATTTGGPYGMCIGAVASVYS 121
 Db 425 VITMESIERKASGKI--GDITSVIDGIAFQTNILALNAAVARAGDEGRFAVYAGRY- 481
 QY 122 SIILFGSSAKNSVAVIDRALSKHDEAIORHAAGAKRDPAS--SAFIOVKQOONL 179
 Db 482 RILORSQAOK--EIKALITESSVAGSQQVQASNAAMDIVSYSDVTTIHSEITNA 540
 QY 180 TDSLSLIANANVPYKFSNFTIGLESRISSGAAATSLSDAKRAVDLILYQOLVYVRETL 239

Db 541 ADEQMR-----GIHEINSNAQOLDTMYQONALVQESTASAA-----LQQAADLFDTV 590

RESULT 6
 AG0893
 aerotaxis receptor protein [imported] - Salmonella enterica subsp. enterica serovar Typh
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 17-May-2002
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Far, S.; Moule, S.; O'Gaora, P. Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0893
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-506 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD07741.1; PID:g16504293; GSPDB:GN00176
 C:Genetics: STY3395
 C:Superfamily: methyl-accepting chemotaxis protein

Query Match 5.1%; Score 118.5; DB 2; Length 506;
 Best Local Similarity 21.2%; Pred. No. 1.3; Indels 115; Gaps 11;
 Matches 68; Conservative 39; Mismatches 99;

QY 24 SDVNSLITKVTALKEASGSNEALELEGKEIOTKPRVGOATKILGVSAGLKLN 83
 Db 262 SQVSVNRGSRRLAKGNNDLNEHRTQYVENQ--ETVTTMNMASVILNSTASADKLIS 320
 QY 84 -----SGDAKTIISGDIYA-----GIATP-----CG 107
 Db 321 MAASSAATOGGEANDFTYKTDIDIAHSTORIGITITTLINDIAFQTNILALNAAVEAARAG 380
 QY 108 PVGIGIGAVAFVSIIISLFGSSAKNSVAVIDRALSKHDEAIORHAAGAKRDEFAESS 167
 Db 381 EOGGFAVVAEVEVHLAS--RSANAANDIRKLIDNSATKVGSGSEOVAAGRTMD----- 433
 QY 168 AFIOVKQOONLTDSDSIIANANVPYKFSNFTIGLESRISSGAA--ATTSLSDAKRAVD 225
 Db 434 ---DIVAQVQVY-----LLARISQSGQEDTGDLSLTRAVD- 468
 QY 226 IILYQOLVYVRETLVDLAILYKGNNAEIVASAVENANRNKELADTLDFLHLIPEQA 285
 Db 469 -----ELNRITOKNAA-----LVEESA 485
 QY 286 LIGA-VYHPIASSETSKAILN 305
 Db 486 QVSAMVHRASRLDPAVTVLH 506

RESULT 7
 S54776
 transcription factor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
 C:Accession: S54776
 R:Evans, R.; Smid, A.; Rudloff, U.; Lottspeltch, F.; Grummt, I. EMBO J. 14, 1248-1256, 1995
 A:Title: Different domains of the murine RNA polymerase I-specific termination facto
 A:Reference number: S54776; MUID:95237206; PMID:7720715
 A:Accession: S54776
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-833 <EVE>
 A:Cross-references: EMBL:X63974; NID:9639818; PIDN:CA58808.1; PID:9639819

Query Match 5.0%; Score 117; DB 2; Length 833;
 Best Local Similarity 22.0%; Pred. No. 3.2;

OY 421 NMFMYMESSASGYIRSWENNPQGHWSIT 450
 DB 421 NMFMYMESSASGYIRSWENNPQGHWSIT 450

RESULT 2

toxin-A - jellyfish (Carybdea alata)

C:Species: Carybdea alata
 C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
 C:Accession: J07372; PC7095
 R:Nagai, H.; Takawa, K.; Nakao, M.; Sakamoto, B.; Crow, G.L.; Nakajima, T.
 Biochem. Biophys. Res. Commun. 275, 589-594, 2000
 A:Title: Isolation and characterization of a novel protein toxin from the Hawaiian box
 A:Reference number: J07372
 A:Accession: J07372

A:Molecule type: mRNA
 A:Residues: 1-463 <NAG>
 A:Cross-references: DDBJ:AB036714
 A:Accession: PC7095
 A:Molecule type: protein
 A:Residues: 158-168;431-463 <NAG>
 C:Comment: This protein, belongs to a class of bioactive proteins with hemolytic activity
 C:Keywords: hemolysis; inflammation; toxin

Query Match 42.7%; Score 992.5; DB 2; Length 463;
 Best Local Similarity 43.5%; Pred. No. 7e-62;
 Matches 197; Conservative 95; Mismatches 148; Indels 13; Gaps 7;

OY 9 LRTVLAITAKHG-----KSDVNSLITVEYALAKAS--GSNEAL-ELALELKEIOT 60
 DB 11 LFLVISTAPPSQARLSRYRRSADAVSTIDIGIQLNDLGTPTKRLKALQGVQAVKK 70
 OY 61 KP-DRVGQATKILGVSALGKLSGDAATKISGCDIYAGTATTEGPGVGMKIGAVAF 119
 DB 71 EATITAKSTIVSGSLSKFKSDPPDVASGCDITIASVATYGGPGIAGVAVSL 130
 OY 120 VSSITSLFTGSSAKNSAAVIDRALSKRDEAIQRHAGAKRDPFASSAFIOYMKOOSNL 179
 DB 131 ISSITSLFSGNSMGSAIKVIDDAFKRYRQDELDVYKAKRFENAVITFVNSVSTENL 190
 OY 180 TPSSDIITANVPYKFSNFIGOLESRISOGAATSTLSAKRAVDITLLYCOLVVRKEL 239
 DB 191 TEVHLDSVDAVRDAFTMLGLVLESIRNGSVSTIDNEMARTINIFLYLDSVARETL 250
 OY 240 LVDLAILYRK--GNAEHVASAVENANRVKELADTLDLHLKILPEQALIGAVYHPISAS 297
 DB 251 LQVILLKVRAGAVDELALSLISLSDOKKEATREIVTLHOMETKYSILGSGTYITDHS 310
 OY 298 ETSKAILNTKTFYGPDPVPR-PIGNRRYKFTNSYWNYSICEAYMGNYMFRGCSNVRNP 356
 DB 311 KAAIILKILKTFEYGPDPARFTFDGLLYMQNRANRYSICKESYAGNMFRGCKOSSYH 370
 OY 357 NRVKMSGFEYTMENSDRRKILTKHOGMGWGLTDEDPGQGHMRFPLRHGKYMS 416
 DB 371 GIRIKLENGYITL-TLSKAMTVYKHAQMGWGTADDEPGGYETFLPLNGEYMYST 429
 OY 417 KRMFMFMYMESSASGYIRSWENNPQGHWSI 449
 DB 430 KKMFDYFYMESSAHGYIRSWHYNPDPOGWKI 462

RESULT 3

toxin-A - Chitropsalmus quadrigatus

C:Species: Chitropsalmus quadrigatus
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: J07805; PC7181

R:Nagai, H.; Takawa, K.; Nakao, M.; Oshiro, N.; Iwanaga, S.; Nakajima, T.
 Biosci. Biotechnol. Biochem. 66, 97-102, 2002
 A:Title: A novel protein toxin from the deadly box jellyfish (see wasp, habu-kurage)
 A:Reference number: J07805; PMID:1186126; MUID:21545932
 A:Accession: J07805

A:Molecule type: mRNA
 A:Residues: 1-462 <NAG>
 A:Cross-references: DDBJ:AB045319
 A:Accession: PC7181
 A:Molecule type: protein
 A:Residues: 432-452;21-54 <NAG>
 C:Comment: This protein is a major proteinaceous toxin in the nematocyst, so that 1

A:Gene: Cqtx-A
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:45-62/Region: amphiphilic alpha-helix-like sequence #status predicted

Query Match 18.2%; Score 423.5; DB 2; Length 462;
 Best Local Similarity 26.3%; Pred. No. 5e-22;
 Matches 120; Conservative 108; Mismatches 171; Indels 57; Gaps 17;

OY 24 SDVNSLITVEYALKEASGSNEALALGLKEIOTKPDVGQATKILGVS---ALG 80
 DB 26 SGLASLAKIDA--KRPSC--KQLFQVANNMQIIEKFSNDEKRAKVMKALGSLSTAVG 81
 OY 81 KLSGDPATKIIISGCDIYAGIATTFEGPGVGMKIGAVAFVSSILSLFTGSSAKNSVAAVT 140
 DB 82 KQSGDPARIASGCDIILVIGISSVLND--FAKSPIFSILSMVGLFSGTKAEESVGSV 139
 OY 141 DRALSKRDEAIQRHAGAKRDPFASSAFIO--VMKQOSNLTSDSITIANVPYKFSNF 199
 DB 140 KRYVQDSOQELQALGYKREYAVSKATLDGVRNETSLSPEVSALGANVPYGVRF 199
 OY 200 IOLESRIISOGAATSTLSDAKRAVDITLLYCOLVVRKELLYDL-AILYRKNAEHVASA 258
 DB 200 IAWVQRIKRRPRTE-SEIKRVLSMLEFTDLCSRLDILDLVOLVATPGHSPRIASG 258
 OY 259 VENANRVKELADTLDLHLKILPEQALIGAVY-HPISASETSKALNTKYFGVDPVR 317
 DB 259 IEVSNLGRREYKVEDELTKNDKETYLELSLYPREHREOSOKIF---KEF---DLAK 312
 OY 318 PIGNRRYK-----FTNSYWNYSICEAYMGNYMFRGCSNVRNPNTRVKMSDGEYT 369
 DB 313 VAYDRLKQDLTGIOVFSLHMPNFTLCSK---DYALITCKPVG-SRLDLKNGEFS 368
 OY 370 MENSRR-----KLYTKHDQMGWGLTDEDPGQGHMRFPLRHK---YVWSSKR 418
 DB 369 IKTQSNPRVCHRYGEYLIFTHDRN-----DLEKFNFPVKLGERKIYLLSSRA 418
 OY 419 WPMFMFMYMESSASGYIRSWENNP-----GPGHWSI 449
 DB 419 SPNKFAIVPKIAKGLDFEYDGIPIQLGYGNOGYFTL 454

RESULT 4

hypothetical protein PX01-79 - Bacillus anthracis virulence plasmid PX01

C:Species: Bacillus anthracis
 C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 16-Feb-2001
 C:Accession: G59100

R:Oknaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Kelm, P.; Koe
 J. Bacteriol. 181, 6509-6515, 1999
 A:Title: Sequence and organization of PX01, the large Bacillus anthracis plasmid har
 A:Reference number: A59091; MUID:99445483; PMID:10515943
 A:Accession: G59100

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1222 <OKI>
 A:Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD32383.1; PID:g4894295
 A:Experimental source: strain Sterne
 C:Note: similar to hypothetical, hydrophobic protein (567 aa), Bacillus firmus (U645
 C:Genetics:
 A:Gene: PX01-79
 A:Genome: plasmid
 C:Superfamily: hypothetical protein PX01-79

Query Match 5.98%; Score 137; DB 2; Length 1222;
 Best Local Similarity 23.0%; Pred. No. 0.21;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 11:14:48 ; Search time 21 Seconds
(without alignments)
2060.023 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322
Sequence: 1 MILKHPWLFYLAITSAKH.....SGYIRSWNNPQGHWSIT 450

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR-73:*

1: PIR1:***
2: PIR2:***
3: PIR3:***
4: PIR4:***

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2317	99.8	450	2 JC7371	toxin - jellyfish
2	992.5	42.7	463	2 JC7372	toxin-A - jellyfish
3	423.5	18.2	462	2 JC7805	toxin-A - Chitropsa
4	137	5.9	1222	2 G59100	hypothetical prote
5	119.5	5.1	595	2 AE0150	probable methyl-ac
6	118.5	5.0	506	2 AG0893	aerotaxis receptor
7	117	5.0	833	2 S54776	transcription fact
8	116.5	5.0	423	1 DEBYMC	malate dehydrogena
9	116	5.0	530	2 C72291	methyl-accepting c
10	116	5.0	1571	2 AC1647	Lactobacillus phag
11	115	5.0	800	2 AG0438	probable Rhs acces
12	112.5	4.8	510	2 C84738	probable kinesin I
13	112.5	4.8	964	2 E71460	probable outer mem
14	112.5	4.8	6713	2 B89921	hypothetical prote
15	111.5	4.8	656	2 A72428	methyl-accepting c
16	111.5	4.8	1336	2 T18288	ABC transport prot
17	110.5	4.8	656	2 E72379	methyl-accepting c
18	110.5	4.8	661	2 G72316	methyl-accepting c
19	110.5	4.8	1024	2 S10056	hemolysin A - Esch
20	110	4.7	810	2 B84185	cytochrome-like pr
21	110	4.7	2297	2 AB2494	hypothetical prote
22	109	4.7	541	2 E83032	probable chemotaxi
23	109	4.7	612	2 B75282	probable acyl-CoA
24	108.5	4.7	659	2 G82365	methyl-accepting c
25	108.5	4.7	1023	1 LEECA	hemolysin A - Esch
26	108	4.7	492	2 A97429	flagk protein prote
27	108	4.7	492	2 AB2647	hook associated pr
28	107.5	4.6	2155	2 C97523	hypothetical prote
29	107.5	4.6	2155	2 AD2742	conserved hypothet

30	106.5	4.6	669	2 AC2624	methyl-accepting c
31	106.5	4.6	692	2 A97406	methyl-accepting c
32	106.5	4.6	1098	2 B70232	hypothetical prote
33	106	4.6	1975	2 B81192	hemagglutinin/hemo
34	105.5	4.5	312	2 C55578	hypothetical prote
35	105	4.5	506	2 E65095	aerotaxis receptor
36	105	4.5	506	2 A85968	aerotaxis sensor r
37	105	4.5	506	2 B91123	aerotaxis sensor r
38	104.5	4.5	532	2 E87343	ABC transporter, A
39	103	4.4	504	2 H87422	methyl-accepting c
40	103	4.4	1456	2 S14005	hypothetical prote
41	102.5	4.4	353	2 A11859	leucine dehydrogen
42	102.5	4.4	460	2 T19111	hypothetical prote
43	102	4.4	394	2 T37853	flagellin, 40K - P
44	102	4.4	1456	2 JQ2284	hypothetical 16S.1
45	101.5	4.4	299	2 B69263	aspartate carboxy

ALIGNMENTS

RESULT 1

toxin - jellyfish (Carybdea rastoni)

C:Species: Carybdea rastoni

C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C:Accession: JC7371; PC7094

R:Nagai, H.; Takawa, K.; Nakao, M.; Ito, E.; Miyake, M.; Noda, M.; Nakajima, T.

Biochem. Biophys. Res. Commun. 275, 582-588, 2000

A:Title: Novel proteinaceous toxins from the box jellyfish (sea wasp) Carybdea rast

A:Reference number: JC7371

A:Accession: JC7371

A:Molecule type: mRNA

A:Residues: 1-450 <NAG>

A:Cross-references: DDBJ:AB015878

A:Accession: PC7094

A:Molecule type: Protein

A:Residues: 39-55;56-70;196-210;250-267;268-279;309-325;363-377;378-382 <NAG>

A:Comment: This protein, a member of bioactive protein, has hemolytic activity.

C:Keywords: hemolysis; inflammation; toxin

Query Match	99.8%	Score 2317;	DB 2;	Length 450;
Best local similarity	99.8%	Pred. No. 1.2e-154;		
Matches	449;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;
QY	1	MILKHPWLFYLAITSAKHGRSDVNSLITRETALEKASGSNPALEGLGEIOT	60	
DB	1	MILKHPWLFYLAITSAKHGRSDVNSLITRETALEKASGSNPALEGLGEIOT	60	
QY	61	KPDHYGOATKILGVSALGKLNISGDAITISGCDIYAGIATTTGGPVGMIGAVASPV	120	
DB	61	KPDHYGOATKILGVSALGKLNISGDAITISGCDIYAGIATTTGGPVGMIGAVASPV	120	
QY	121	SSILSFTGSSAKNSVAVIDRALSKHRDEAIORHAAKRDFASSAFIOYMOOSMLT	180	
DB	121	SSILSFTGSSAKNSVAVIDRALSKHRDEAIORHAAKRDFASSAFIOYMOOSMLT	180	
QY	181	DSLSIIAANPVYKFSNFIQLESRIISOGAATTSLSAKRAVDITLLYCOLVYKRETL	240	
DB	181	DSLSIIAANPVYKFSNFIQLESRIISOGAATTSLSAKRAVDITLLYCOLVYKRETL	240	
QY	241	VDLALYRKGAHEVASAVENANRNYKELADTDLFLKLLPEOALIGAVYHPISASETS	300	
DB	241	VDLALYRKGAHEVASAVENANRNYKELADTDLFLKLLPEOALIGAVYHPISASETS	300	
QY	301	KAILNYTYFGVPDVPDRIGNRKRYKFTNSYNTYSICSEAYMGVYMGCSNVNPNIRV	360	
DB	301	KAILNYTYFGVPDVPDRIGNRKRYKFTNSYNTYSICSEAYMGVYMGCSNVNPNIRV	360	
QY	361	SKMSGCEYTMESDRRKLYITRKHOOGMGCTLDEPQGGHRRFPLHGGYMWSSRRMP	420	
DB	361	SKMSGCEYTMESDRRKLYITRKHOOGMGCTLDEPQGGHRRFPLHGGYMWSSRRMP	420	

Best Local Similarity 21.9%; Pred. No. 39;
Matches 75; Conservative 40; Mismatches 127; Indels 100; Gaps 16;

```

QY 42 GSENALELEGLKEIOTKPR--VGQATKILGSGSALG--KL--NSGDATKI--I 91
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 70 GISEA-----DGGKGGANARGDKSIAGIDIAQALGSGSIAIGDNKIYHNSNNANIGAKA 124
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 92 SGCLDIVAGIATTFGPGVMGIGA-----VAFVSSILSLFTGSSAKNSVAAYIDRA 143
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 125 SGNESIAIGGDVLAGSHASIAIGSDLYLKKEVVOQISELPIIRGQKALNDIYQLADTN 184
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 144 LSKHDEAIORHAGAKRDFAESAFIQVMKQOSNLTDSLSITIANVPYKESNFIGOL 203
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 185 LQYRRTHAQGHASTA-----VGAMSTAKGHFSNAFGTR 218
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 204 ESRIQGAATTSLSDAKRAVDLILYCOLVVRRETLVLAALYRKNAEHV---ASAVE 260
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 219 ATAECTYSLAVGLTATAKAASIAY-----GSNAQAIGFAATAVG 258
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 261 NANRYNKLADTLDLKLIPQALIGAVYHPISASETSKALINYYTYFGVDPVPRPIG 320
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 259 GSTOVNLRGI-ALGFGSQVLQKNDVNAA-----NVRAY--APDDNQPID 301
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 321 NRRYK--FTNSYWNYYTSCSEAYMGNYMFRGCSNVRNPDIRV 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 302 N-RYKATFEKNGATDVEFI-----GN--SNGNDSIRKRIINV 334
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Search completed: May 19, 2003, 11:28:04
Job time : 69 secs

```
RESULT 13
US-09-815-242-12611
; Sequence 12611, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELTRA-011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12611
; LENGTH: 3158
; TYPE: PR
; ORGANISM: Staphylococcus aureus
US-09-815-242-12611

Query Match
Best Local Similarity 4.3%; Score 100; DB 10; Length 3158;
Matches 66; Conservative 65; Mismatches 147; Indels 54; Gaps 14;

QY 17 SAKHGRSDVNSLTFTVETLAKFASGSE-----ALLEALEGKGEIQTK-PDRVG 66
DB 1054 TGGNANQPEVQAITRYVQTLTQALGDHNLQVAKTNATQALDALTSLNDPQTKALKDQVT 1113
QY 67 QATKILGVSAGLKL--NSGDATKTIISGCLDIIVAGIATPFGG-----PYGMIGAV 116
DB 1114 AATLV-----TAVHQTIEQNANTLNQAMHGLRESIODNAATKANSKYINEDQPEQANTDOA 1168
QY 117 ASFVSSILSFTG---SSAKNSVAAYIDRA-LSKHDEAIOHRAAGAKRDPASSAFIOY 172
DB 1169 VQANSTIINQATLNDNNALNQAAATVNTKALHGDVYKQNDKAKQVQSOLAYLNNA 1228
QY 173 MK-QQSNLSDSLSTIAANPVYKFSNFICQLSRSISQGAATISLSAKAAVDFIL----- 227
DB 1229 QKHEMTLLIDSETTRIAVKODLTE-AQALDQUNMTLQOQSIND--KDATBASSAYVAEP 1284
QY 228 ----LYCOLVVMRETLIVDL-ALLYRKNAEHVAASAVENANR-----VNKELADT 273
DB 1285 NKQAVDEAVQNAESIAGLNPTINKGNVSSATQAVTTSKNGLDGVERLAQDKQTAGNS 1344
QY 274 LDFLHLKLP--EQALIGAVYHPISASETSKAI 303
DB 1345 LNHLDQLTPAQOQALQEMQINNATTBPKVAEII 1376

RESULT 14
US-10-125-692-13
; Sequence 13, Application US/10125692
; Publication No. US2003004429A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Agerem, Alan
; APPLICANT: Hayashi, Fumitaka
; APPLICANT: Smith, Kelly D.
; APPLICANT: Underhill, David M.
; APPLICANT: Ozinsky, Adrian
; TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
; FILE REFERENCE: P-15 5155
; CURRENT APPLICATION NUMBER: US/10/125,692
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/285,477
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 379
; TYPE: PR
; ORGANISM: V. cholerae
US-10-125-692-13

Query Match
Best Local Similarity 4.2%; Score 98.5; DB 9; Length 379;
Matches 66; Conservative 37; Mismatches 85; Indels 131; Gaps 12;

QY 13 LAITSAKHGRSDVNSLTFTV-ETALKEASGSEDA-LEALEGKGEIQTKPDRVQCAT- 69
DB 72 ISTAQTBEGAMNVESTILQMRDLQLOANGTNSASEROLNBSVAQLODELRIAEFTTS 131
QY 70 ----KIL-GSVGSLAKLNSGDATKIISG----- 93
DB 132 FGGRKILNGSFGASFOIGSSGEALIMGLTSVRADFRMGQSFIAEQPKTKEMGVPT 191
QY 94 -----CLDIVA----- 99
DB 192 ARDLKEFTTKDGEAVVLDIARDGDDIEELATYINGQTDLFKASVDOEGKQIFVAEPN 251
QY 100 -----GIATTFGGPVGMIGAVASFVSSILSFTGSSAKNSVAAYIDRALSKRDE 150
DB 252 IEGNFNISGLATELGLNGSPGVKTYVODI-----ITSVGGSONAVGITDAL-KYVDS 305
QY 151 AIORHAAGAKRDPASSAFIOVAKQOQSNLTDSDLSITIAANVPYK-----FSNFIGOL 203
DB 306 --QRADLGAKQNLKSHI-----SNLSIQENVEVSKSHIKDTPRAKETTLQ 350
QY 204 -ESRISQGAATTSLSDAKR 221
DB 351 TKSQILQOAGTSIIAQAKQ 369

RESULT 15
US-09-813-214A-9
; Sequence 9, Application US/09813214A
; Patent No. US20020177200A1
; GENERAL INFORMATION:
; APPLICANT: Tucker, Kenneth
; APPLICANT: Piosilla, Laura
; TITLE OF INVENTION: MORAXELLA CATARRHALIS OUTER MEMBRANE PROTEIN-106 POLYPEPTIDE,
; FILE REFERENCE: 7969-089-999
; CURRENT APPLICATION NUMBER: US/09/813,214A
; CURRENT FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 08/968,685
; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 2122
; TYPE: PR
; ORGANISM: Moraxella catarrhalis
US-09-813-214A-9

Query Match
Best Local Similarity 4.2%; Score 97; DB 9; Length 2122;
```

PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 5639
 LENGTH: 2086
 TYPE: PRF
 ORGANISM: Staphylococcus aureus
 US-09-815-242-5639

Query Match 4.5%; Score 104.5; DB 10; Length 2086;
 Best Local Similarity 20.7%; Pred. No. 8.9;
 Matches 79; Conservative 59; Mismatches 134; Indels 109; Gaps 18;

QY 12 VLATSAKGRSDVNSLITVETALKEASGN-----EALALEG-----L 54
 DB 1095 VINAASNPMDANAIQATVSTKNALDGTNLTQAQOTAINAIDGATNLNKAQKDL 1154
 QY 55 KEIOTKPRVGOATKI---LGSVSGALGKNSG---DATKISGLDIYAGIATTEGG 107
 DB 1155 KAOV-TSAORVANVTISQGTANELNLTAMGLOHGDIDENATKOTOKYRAEOSKRAYD- 1212
 QY 108 PYGMCIGAVASFVSSILSTFGSSAKNSVAAYIDRALSK-----HRDEAIQRHAGA 159
 DB 1213 ---QAVAA-AKAIINMOTGS---NSDKAAVDRLQOVSTSKALNDALAEKAA 1262
 QY 160 KRDF-----AESAF-----IOVMKOOSMLTDSIITANV---PYVK 195
 DB 1263 KQNLGTNLNHTNAORTALEGOINQATTVDGVTVTKNANTLGGAMNSLOGSINDKQATLR 1322
 QY 196 FSNFISGLESRIIS-----OGATISLSDAKRAVDLILYCOLVARETL- 239
 DB 1323 NQNTYDADESKRNATQAVTALEGLNKTGNTSADVDNLTNTYRAKALNGAENLR 1382
 QY 240 -----LVDLAIIRKGNAEHVASAVENANRNV--KELADLDFLHKLIPEDA 285
 DB 1383 NTKTSATNTINGLPMLTOL-QKDNLKHVEQAQNVAGVGVAD-KGNTLN-----T 1431
 QY 286 LIGAVYHPIASSETSKALINY 306
 DB 1432 AMGALRTSIQNDNTTKTSQNT 1452

RESULT 11
 US-10-125-692-14
 Sequence 14, Application US/10125692
 Publication No. US2003004429A1
 GENERAL INFORMATION:
 APPLICANT: Aderem, Alan
 APPLICANT: Hayashi, Fumitaka
 APPLICANT: Smith, Kelly D.
 APPLICANT: Underhill, David M.
 TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
 TITLE OF INVENTION: of Use
 FILE REFERENCE: P-1S 5155
 CURRENT APPLICATION NUMBER: US/10/125,692
 PRIOR FILING DATE: 2002-04-17
 PRIOR APPLICATION NUMBER: US 60/285,477
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 43
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 14
 LENGTH: 394
 TYPE: PRF
 ORGANISM: P. aeruginosa
 US-10-125-692-14

Query Match 4.4%; Score 102; DB 9; Length 394;

Best Local Similarity 21.3%; Pred. No. 1.4;
 Matches 71; Conservative 39; Mismatches 92; Indels 132; Gaps 13;

QY 11 IYLAITSAKHGRSDVNSLITVETALKEASGN-EALALEGLKGTOTPRDRVGOAT 69
 DB 72 ISLAQT-ABGALQOSTNIIQRRMDLSLOSANSNSDSERTALNGEAKOLKEIDRISNT 130
 QY 70 -----KIL-GSVGSLGKNSGDATKII-----SG 93
 DB 131 TFGGRKLDGSPGVASFQVGS-AHEITISVIGIDENSAESLNGTYERKADGGAVTAAATAG 189
 QY 94 CEDIYAGIATTEGPGVM----- 111
 DB 190 TVDIALGI--TGSANVNVVDKMGNETAQAATAAVNDANVGIGAFSDGDTISYVK 247
 QY 112 ---GIGAVASFVSSILSTFGSSA-----KNSVAVID---RALSXH 147
 DB 248 AKGDSGAIITSAVSGVIAIDTGTGVTGAAGVAPATAFAKTNDRVAKIDISTAKLSRR 307
 QY 148 ---RDEAIQRHAGAKRDFEASAFLOVKKOOSN-----LTDSLSITIANV 193
 DB 308 AGDRTTALKQIDASVPTSYAONRFDNTNLTNLTGENVSAARGRIEDTFAETANTL-- 365
 QY 194 YKFSNFIGLESRIISQGAATTSLSDAKRAVDLIL 227
 DB 366 -----TKQVLAQAGTALIAQANQLPOSVL 390

RESULT 12
 US-09-820-843A-50
 Sequence 50, Application US/09820843A
 Publication No. US2003003963A1
 GENERAL INFORMATION:
 APPLICANT: Council of Scientific and Industrial Research
 TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PF
 FILE REFERENCE: 063915
 CURRENT APPLICATION NUMBER: US/09/820,843A
 CURRENT FILING DATE: 2001-03-30
 NUMBER OF SEQ ID NOS: 118
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 50
 LENGTH: 444
 TYPE: PRF
 ORGANISM: C. pneumoniae CWL029
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: CT579 hypothetical protein
 NAME/KEY: misc_feature
 OTHER INFORMATION: g14377120
 US-09-820-843A-50

Query Match 4.3%; Score 100; DB 9; Length 444;
 Best Local Similarity 24.7%; Pred. No. 2.5;
 Matches 54; Conservative 34; Mismatches 89; Indels 42; Gaps 8;

QY 28 SILTKVETALKEASGNEALALEGLKGTOTPRDRVGOATKILGSVSGALGKNSGDA 87
 DB 143 ALLQVNTLMAANAQ-----ESWKASQSONAIRSOFESAPRIGETIKRQANHQA 193
 QY 88 T-----KIISGLDIYAGIATTEGPGVMGI-----GAVASFVSSILSTFGSSAKN 134
 DB 194 SATPAQKOSLISIVNG-----FTVSVGAGIFSAKAKATSLKASAKETGASAG 248
 QY 135 SYAAVIDRALSKHDEAIQRHAGAK-RDPAESSAFLOVKKOOSNLTDSDLSTIANV 193
 DB 249 GAAS---KALTSASSVOQTMASTAKATTAASSAGSAATKAANLTD-DMAAASAKMAS 304
 QY 194 YKFSNFIGOL-----ESRISQGAATTSLSDAKRA 222
 DB 305 DGASKASGGLFGEVTLNKPNNSEKYSRGMNVYKQGANVA 343

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5883
LENGTH: 837
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5883

Query Match 4.5%; Score 104.5; DB 10; Length 837;
Best Local Similarity 19.1%; Pred. No. 2.5;
Matches 75; Conservative 71; Mismatches 147; Indels 99; Gaps 17;

QY 16 TSAHGRSDVNSLTETKETAKEASGN-----EALALEGLKGEIQ----- 59
DB 301 TDADAKRNATYNAVTOAEQILNKAGPNTSKDGVETALENQRANKNELNGNQNANAKT 360
QY 60 TKPRVGOATKILGSVSGALGKLSGDAATKIIISGLDIYAGIATTFGGPVGIGAVASF 119
DB 361 TAKNALNLTSINNQRKALSKQIEGAT-----VAGVNO-----VSTTASE 402
QY 120 VSSILS-LFTGSSAKNSAAVYD-----RALSKHDEAIO--RHAAGAKRDPFAESSAFIQ 171
DB 403 LNTAMSNLQNGINDEAAKALNGTONLEKAKQHANTAIIDGLSHLTNAQKE-----ALKQ 457
QY 172 VMKQOONLTDSDLSIIANPVYKFSNFIQLESRIISOGAAT-----TSLSDA-----KRAVD 224
DB 458 LVQOSTYAEAGNQRKAN-----NDAMMDKLRQSIADNATTKONONTDASQNKADAYN 513
QY 225 FILYQCQVLMRETL-LVDLAILYR-----KGAHEVASAVENA 262
DB 514 NAVTAAQGIIDQTSPTLDPTVINOAGOVSTTKNALNGENLEAKQAQASQSLSDNL 573
QY 263 NRVKLEADPLDPLHLK-----IPEQA-----LIGAYHPIASSETSKALINTKRYGV 312
DB 574 NNAQQTVDIDINGAHYVDEANQIKONQONMTANGNLKQALADDAKATVNF----- 628
QY 313 PDVPRIGNRKRYKFTNSYNTSICSEAYMGN 344
DB 629 -DADQA---KQAYNTAVTNAENITSKANGN 656

RESULT 9
US-09-815-242-13080

Sequence 13080, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13080
LENGTH: 875
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-13080

Query Match 4.5%; Score 104.5; DB 10; Length 875;
Best Local Similarity 19.1%; Pred. No. 2.7;
Matches 75; Conservative 71; Mismatches 147; Indels 99; Gaps 17;

QY 16 TSAHGRSDVNSLTETKETAKEASGN-----EALALEGLKGEIQ----- 59
DB 338 TDADAKRNATYNAVTOAEQILNKAGPNTSKDGVETALENQRANKNELNGNQNANAKT 397
QY 60 TKPRVGOATKILGSVSGALGKLSGDAATKIIISGLDIYAGIATTFGGPVGIGAVASF 119
DB 398 TAKNALNLTSINNQRKALSKQIEGAT-----VAGVNO-----VSTTASE 439
QY 120 VSSILS-LFTGSSAKNSAAVYD-----RALSKHDEAIO--RHAAGAKRDPFAESSAFIQ 171
DB 440 LNTAMSNLQNGINDEAAKALNGTONLEKAKQHANTAIIDGLSHLTNAQKE-----ALKQ 494
QY 172 VMKQOONLTDSDLSIIANPVYKFSNFIQLESRIISOGAAT-----TSLSDA-----KRAVD 224
DB 495 LVQOSTYAEAGNQRKAN-----NDAMMDKLRQSIADNATTKONONTDASQNKADAYN 550
QY 225 FILYQCQVLMRETL-LVDLAILYR-----KGAHEVASAVENA 262
DB 551 NAVTAAQGIIDQTSPTLDPTVINOAGOVSTTKNALNGENLEAKQAQASQSLSDNL 610
QY 263 NRVKLEADPLDPLHLK-----IPEQA-----LIGAYHPIASSETSKALINTKRYGV 312
DB 611 NNAQQTVDIDINGAHYVDEANQIKONQONMTANGNLKQALADDAKATVNF----- 665
QY 313 PDVPRIGNRKRYKFTNSYNTSICSEAYMGN 344
DB 666 -DADQA---KQAYNTAVTNAENITSKANGN 693

RESULT 10
US-09-815-242-5639

Sequence 5639, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23

Publication No. US20030035809A1
GENERAL INFORMATION:
APPLICANT: GEORGE, LISLE W
APPLICANT: ANGELOS, JOHN A
APPLICANT: HESS, JOHN F
TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
TITLE OF INVENTION: BOVIS INFECTIONS
FILE REFERENCE: 481.06
CURRENT APPLICATION NUMBER: US/09/884,696
CURRENT FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 5
LENGTH: 1023
TYPE: PRT
ORGANISM: Escherichia coli
US-09-884-696-5

Query Match
Best Local Similarity 4.7%; Score 108.5; DB 9; Length 1023;
Matches 82; Conservative 65; Mismatches 163; Indels 113; Gaps 17;

QY 24 SDVNSL---LTKVETALKEAGSNEALEGLKEIQTGP--DRVQATKILGSVGA 78
DB 199 NNWNSFQQLNKLGSLVNT-----KHLNGVGNKQLNPMLNDIGAGIDTVSGILSA 250
QY 79 LGR---LNSGDA---TIIISGCLDIVAGIATTFGPGYMGIGAVAFVSSTLSFTGSSA 132
DB 251 ISASFILSNADADGTGKAAAG---VELTFTKVLGNVGKIGISQITIIORAAOGISTSA 305
QY 133 KNSVAAYIDALSK-----HRDEALORHAAGAKRDEAFSSAPITQVKKQOSNLTD 182
DB 306 AGLASVYTLALISPLSLIADKREKRAKIEEVSQRKFKIGYDDSLAFAHFKETGAIDA 365
QY 183 DLSIAANVYKFSNFIGOLESRISOGAATTSLSDAKRAVDFTLLCOLVAMEETLVD 242
DB 366 SLT-----RISTYTLASVSSGIS-AAATTSLVGAPVSA-----LVGAVTGITSG 407
QY 243 LALLYKRGNAEHYASAV-----ENANFVNKELADLTDLFLKLIPEQALLGAVYHPISAS 297
DB 408 ILBASQAMFEHVASKADYIAEWEKKHGNFYENGYDARHAAFL-----453
QY 298 ETSKAILNTKYFGVDPVPRPIGNRRKFTNSWNTYSTICEAIVMGVNYMRCGSNVNPN 357
DB 454 DNKILISQYKREYSV-----ERSVLTIOQHMDTL-----482
QY 358 IRVSKMSDGEFTYMSNDRR---KLTYTKHDQGWGCTIDPDGDGHRFIPLRHGKMY 414
DB 483 --IGELAG--VTRNGDKTLGSKSYIDYEEG---KRLKKRPDPKQVDPDLKGNIDLS 534
QY 415 SSK 417
DB 535 DSK 537

RESULT 7
US-09-817-514A-8
Sequence 8, Application US/09817514A
Patent No. US2002007847B1
GENERAL INFORMATION:
APPLICANT: French-Constant, Richard
APPLICANT: Bowen, David
APPLICANT: Rochelleau, Thomas
APPLICANT: Waterfield, Nicholas
TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
FILE REFERENCE: 61645
CURRENT APPLICATION NUMBER: US/09/817,514A
CURRENT FILING DATE: 2000-03-26
PRIOR APPLICATION NUMBER: US 60/191806
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0

SEQ ID NO 8
LENGTH: 2504
TYPE: PRT
ORGANISM: Photorhabdus luminescens
US-09-817-514A-8

Query Match
Best Local Similarity 4.5%; Score 105.5; DB 10; Length 2504;
Matches 100; Conservative 85; Mismatches 184; Indels 151; Gaps 27;

QY 11 IVLAITSKHKGRS---DVNSLTKVETALKEAGSNEALEAL-----EG 53
DB 682 LTAISLTHGKESLIGDILKRAMAPCTSAHLT-SEFVAYDLMLTDOQPAQITVDG 740
QY 54 LKGEIQRKPDVY-----GQATKILGSVGSALGKINSGDATKIISGCLDIVAGIATTEGCP 108
DB 741 FWEEOVTPPSLKVITPAQVLAQSLAYRRIG-LETELISLTVQSSILVAGKSTILDHGL 799
QY 109 VGM-----GIGAVAFVSSTLSLFTGSSAKNSVNAVIDRALSKHDEALORHAA 157
DB 800 LITMALGCFHTWVNGLOHSLIAL-----KQALVTVDVAQAMNKEESLDMWA 851
QY 158 G-ARRD-----FAESSAFIQVKKQOSNLTDSDLSIAANVPYKFS---NFIGOLESRI 207
DB 852 NOVEKDLTKLTSWTOIDAILQWLMSSALAVSPIDL--AGMMALKYGDHNYAAM-----904
QY 208 SQAATTSLSD---ARRAVD--FILLYCOLVYRFTLVLDLALYKRGNAEHYASAVEN 261
DB 905 -QAAALALMDHANQAKKIDETFSKALCYII---NAVDSAGVDRNGILTYLLIDN 960
QY 262 ---ANVFNKELADLTDLFLKLIPEQALLGAVYHPISASSETSAIINTKYRGVDP 314
DB 961 QVSADYTTSLAIALIAGIQLYVNRALNRDGOIA-----SDVSTQFTDMERY-----1009
QY 315 VPRPIGNRRYKFTNSWNTYSTICEA--YMGVNYMFCGSNVNPNIRV--SKMSDGEFTYM 370
DB 1010 ---NKR-----STWAGVSELVYVENV-----DPTQRIQGTKMDALLOS 1048
QY 371 EN-----SPRKLTYTKHDQGWGCTL---DEPDGDGHRFIP-----407
DB 1049 INOSQNLADTVEPAFTYTLTSEFQVANLKVISAHYHNVNVDGLTYFGIDQAAPGYW 1108
QY 408 -----RHKYVSSKRPNFMFMESSASGYIRSMEN 439
DB 1109 RSYDHSKCENGKP--AANAMGEV-----NKITCAVNPWKN 1141

RESULT 8
US-09-815-242-5883
Sequence 5883, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELTRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23

Db 1863 NEAKNTAKOQVATMMSHLTDAOKANLTS-----QIESG---TTVAGYOGI----- 1903

Qy 126 LETGSSAKNSVAVIDRALSKHDEAIORHAAGAKRDEFAESSAFIQVMKQOOSLTDSDS 185

Db 1904 -----QANAGTLDQANQLOQSASJASDATKSSSEDYODANADLONANVADA-VTAAGI 1954

Qy 186 IIAANVPYKFSNFIGLESRSISOGAATTSLSAKRAVDIFILLYCOLVYMRRELLVDLAI 245

Db 1955 ISATNPEKN-PDTINOKASOVN-SAKSALNGDEK-----LAAKQAKSDIGR 2001

Qy 246 LYRKGNAEHVASAVENANVKNELADTLDFLKLIPEDALIGAVYHPISASETSKALIN 305

Db 2002 LIDLNNAOQTANAFAEVDQAPN-LAAVTA-AKKKATSLMTAMGNLHALAKXONTKRSV 2058

Qy 306 YF 307

Db 2059 YF 2060

RESULT 4

US-09-841-132-177

; Sequence 177, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yaser A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 177

; LENGTH: 964

; TYPE: PRF

; ORGANISM: Chlamydia

US-09-841-132-177

Query Match 4.8%; Score 112.5; DB 10; Length 964;

Best Local Similarity 21.3%; Pred. No. 0.64;

Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

Qy 21 GKRSDV-----NSLITVETALKEASGSEBALEALGELGKEIOTKPDYRGQATKIIGSVG 76

Db 254 GKRGNIVFYNNRCFKNVEFASSEAS-----DGAIKVTRLDVYGNRGRIFFS-- 301

Qy 77 SALGRKNSGDATRKIIISGCLDIYAGIATTFGGPYGMGIGAVASPVSSILSFTGSSAKNSV 136

Db 302 -----DNITKNYGGAI--YAPVYLVLDGPTTFINNINANKG 336

Qy 137 AAVIDRALSKHDEAIORHA-----AGAKR-----DFAESSAFIQV 172

Db 337 GAIIYDGTNSKRISA-DRAHIIFNENIYVNTVANGTSISANPRRNATIVASSGELL 395

Qy 173 MKQOOS-NLTDSD-LSIIAANVPYKFSNFIGLESRSISOGAATTSLSDAKRAVD----- 224

Db 396 GAGSSONLIFYDPIEVSNAQVSF-SFNKEADQTSYVFSGATVNSADFHORNLQOTKTPAP 454

Qy 225 -----FILL--YCOLVYMRRELL--LVLDL--AII--YRKGNAEHVASAVENANVKNKE 268

Db 455 LTLNSGFLCIEDHAQLTVNRFTOTGGVSLGNCVAVLSCTKNGTGD-----SASASITLKH 510

Qy 269 LAADTLDFLKLIPEDALIGAVYHPISASETSKALINLT-----KYGVDPVPRPI---G 320

Db 511 IGLN-----LSIILKSGAEI-----PLWVEPTNNSNNYTAADTAATFSLSDVKLSLDDYG 561

Qy 321 NRRYKFTNSYMTYISICSEAYMGNYMFRGCSNVRNINIRVSKMSDGFYTMENSDRRKLYI 380

Db 562 NSPYESTDL--THALSSO-----PMLTISEASDNQLOSENIDFSGLVN 602

Qy 381 TKHD-OG-----WGWG-TLDEDDPGDOGHMR-----FIPL-RHGKYNV 414

Db 603 PHYGMGLTWGAKTQDEPPASSATITPDOKANRHRFTLLTLWLPAGVPSPKHNSPLI 662

Qy 415 SSKRPNMFMWESSASGYIRSWENPNPGQGHST 450

Db 663 ANTLMGNMLATESLKN-----SALPLPSGHPFWGIT 694

RESULT 5

US-09-841-132-191

; Sequence 191, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yaser A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 191

; LENGTH: 977

; TYPE: PRF

; ORGANISM: Chlamydia

US-09-841-132-191

Query Match 4.8%; Score 112.5; DB 10; Length 977;

Best Local Similarity 21.3%; Pred. No. 0.65;

Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

Qy 21 GKRSDV-----NSLITVETALKEASGSEBALEALGELGKEIOTKPDYRGQATKIIGSVG 76

Db 267 GKRGNIVFYNNRCFKNVEFASSEAS-----DGAIKVTRLDVYGNRGRIFFS-- 314

Qy 77 SALGRKNSGDATRKIIISGCLDIYAGIATTFGGPYGMGIGAVASPVSSILSFTGSSAKNSV 136

Db 315 -----DNITKNYGGAI--YAPVYLVLDGPTTFINNINANKG 349

Qy 137 AAVIDRALSKHDEAIORHA-----AGAKR-----DFAESSAFIQV 172

Db 350 GAIIYDGTNSKRISA-DRAHIIFNENIYVNTVANGTSISANPRRNATIVASSGELL 408

Qy 173 MKQOOS-NLTDSD-LSIIAANVPYKFSNFIGLESRSISOGAATTSLSDAKRAVD----- 224

Db 409 GAGSSONLIFYDPIEVSNAQVSF-SFNKEADQTSYVFSGATVNSADFHORNLQOTKTPAP 467

Qy 225 -----FILL--YCOLVYMRRELL--LVLDL--AII--YRKGNAEHVASAVENANVKNKE 268

Db 468 LTLNSGFLCIEDHAQLTVNRFTOTGGVSLGNCVAVLSCTKNGTGD-----SASASITLKH 523

Qy 269 LAADTLDFLKLIPEDALIGAVYHPISASETSKALINLT-----KYGVDPVPRPI---G 320

Db 524 IGLN-----LSIILKSGAEI-----PLWVEPTNNSNNYTAADTAATFSLSDVKLSLDDYG 574

Qy 321 NRRYKFTNSYMTYISICSEAYMGNYMFRGCSNVRNINIRVSKMSDGFYTMENSDRRKLYI 380

Db 575 NSPYESTDL--THALSSO-----PMLTISEASDNQLOSENIDFSGLVN 615

Qy 381 TKHD-OG-----WGWG-TLDEDDPGDOGHMR-----FIPL-RHGKYNV 414

Db 616 PHYGMGLTWGAKTQDEPPASSATITPDOKANRHRFTLLTLWLPAGVPSPKHNSPLI 675

Qy 415 SSKRPNMFMWESSASGYIRSWENPNPGQGHST 450

Db 676 ANTLMGNMLATESLKN-----SALPLPSGHPFWGIT 707

RESULT 6

US-09-841-696-5

; Sequence 5, Application US/09884696

Matches 64; Conservative 58; Mismatches 121; Indels 59; Gaps 12;

QY 16 TSAHGRSDVNSLTITKVEALKEASGSNE--AALE-----ALBGLKEIQTKPDV 65
Db 1029 TDADREKOTAYNDVTAKTLLDKTAGSNDKKAVEQALQRVNTAKTALND-----ERL 1093

QY 66 GOATKILGSVSGALGKNSGDATKTIISGCLDIYAGIATFTGGPVGMGIGAVASFVSSILS 125
Db 1084 NEAKNTAKQOOLATWSHLLTNOKANLTS--QIERG--TTVAGVQGI-----1124

QY 126 LFTGSSAKNSVAIVIDRALSKHDEAIQRAHAGAKRDEFASSAFIYVMOQSNITDSDLS 185
Db 1125 -----QANACTLDOAMQRLQSTASKCATKTSSEDIYQANADLQNAVYHA--VTNAEGI 1175

QY 186 ITAANPVYKESNFICQLESRIISGAAATSLSDAKRAVDLILLYCOLVNRFTLLVDLAI 245
Db 1176 ISATNPEMN-PDTINOKASQVN--SAKSALNGDEK-----LAAKQSAKTIDGH 1222

QY 246 LYKGAHEHVASAVENANRNKRELADTDLFLKLIPEQALIGAVYHPISASETSKAILN 305
Db 1223 LFDLNNAQOTSAVAEVDQAPN--LAAVTA--AKNKATSLNTAMGNLKHALLAEKDNTRKRSVN 1279

QY 306 YT 307
Db 1280 YT 1281

RESULT 2
US-09-815-242-12610
; Sequence 12610, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12610
; LENGTH: 5795
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12610

Query Match 5.1%; Score 118.5; DB 10; Length 5795;
Best Local Similarity 20.9%; Pred. No. 2.4; Indels 59; Gaps 12;
Matches 63; Conservative 59; Mismatches 121; Indels 59; Gaps 12;

QY 16 TSAHGRSDVNSLTITKVEALKEASGSNE--AALE-----ALBGLKEIQTKPDV 65
Db 1029 TDADREKOTAYNDVTAKTLLDKTAGSNDKKAVEQALQRVNTAKTALND-----ERL 1093

Db 4486 TDADREKOTAYNDVTAKTLLDKTAGSNDKKAVEQALQRVNTAKTALNGDA-----RL 4540

QY 66 GOATKILGSVSGALGKNSGDATKTIISGCLDIYAGIATFTGGPVGMGIGAVASFVSSILS 125
Db 4541 NEAKNTAKQOOLATWSHLLTNOKANLTS--QIERG--TTVAGVQGI-----4581

QY 126 LFTGSSAKNSVAIVIDRALSKHDEAIQRAHAGAKRDEFASSAFIYVMOQSNITDSDLS 185
Db 4582 -----QANACTLDOAMQRLQSTASKCATKTSSEDIYQANADLQNAVYHA--VSDAEGI 4632

QY 186 ITAANPVYKESNFICQLESRIISGAAATSLSDAKRAVDLILLYCOLVNRFTLLVDLAI 245
Db 4633 ISATNPEMN-PDTINOKASQVN--SAKSALNGDEK-----LAAKQSAKTIDGH 4679

QY 246 LYKGAHEHVASAVENANRNKRELADTDLFLKLIPEQALIGAVYHPISASETSKAILN 305
Db 4680 LFDLNNAQOTSAVAEVDQAPN--LAAVTA--SAKNKATSLNTAMGNLKHALLAEKDNTRKRSVN 4736

QY 306 YT 307
Db 4737 YT 4738

RESULT 3
US-09-815-242-12996
; Sequence 12996, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12996
; LENGTH: 6281
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match 5.1%; Score 118.5; DB 10; Length 6281;
Best Local Similarity 21.2%; Pred. No. 2.7; Indels 59; Gaps 12;
Matches 64; Conservative 58; Mismatches 121; Indels 59; Gaps 12;

QY 16 TSAHGRSDVNSLTITKVEALKEASGSNE--AALE-----ALBGLKEIQTKPDV 65
Db 1808 TDADREKOTAYNDVTAKTLLDKTAGSNDKKAVEQALQRVNTAKTALND-----ERL 1862

QY 66 GOATKILGSVSGALGKNSGDATKTIISGCLDIYAGIATFTGGPVGMGIGAVASFVSSILS 125
Db 4486 TDADREKOTAYNDVTAKTLLDKTAGSNDKKAVEQALQRVNTAKTALNGDA-----RL 4540

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 11:17:28 ; Search time 57 Seconds
(without alignments)
761.454 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2332
Sequence: 1 MILKHPWLFYLAITSKHK.....SGYISWENNEPQCHMSIT 450

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/PCCT_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118.5	5.1	2434	10 US-09-815-242-5835	Sequence 5835, Ap
2	118.5	5.1	5795	10 US-09-815-242-12610	Sequence 12610, A
3	118.5	5.1	6281	10 US-09-815-242-12996	Sequence 12996, A
4	112.5	4.8	964	10 US-09-841-132-177	Sequence 177, App
5	112.5	4.8	977	10 US-09-841-132-191	Sequence 191, App
6	108.5	4.7	1023	9 US-09-884-696-5	Sequence 5, Appl
7	105.5	4.5	2504	10 US-09-817-514A-8	Sequence 8, Appl
8	104.5	4.5	837	10 US-09-815-242-5883	Sequence 5883, Ap
9	104.5	4.5	875	10 US-09-815-242-13080	Sequence 13080, A
10	104.5	4.5	2086	10 US-09-815-242-5639	Sequence 5639, Ap
11	102	4.4	394	9 US-09-820-843A-50	Sequence 14, Appl
12	100	4.3	444	9 US-09-815-242-5194	Sequence 12611, A
13	100	4.3	3158	10 US-09-815-242-12611	Sequence 13, Appl
14	98.5	4.2	379	9 US-10-125-692-13	Sequence 9, Appl
15	97	4.2	2122	9 US-09-813-214A-9	Sequence 459, App
16	95	4.1	434	9 US-09-892-877-459	Sequence 12888, A
17	94	4.0	722	10 US-09-815-242-12888	Sequence 5803, Ap
18	94	4.0	991	10 US-09-815-242-5803	Sequence 5908, Ap
19	94	4.0	1215	10 US-09-815-242-5908	

20	94	4.0	1269	10 US-09-815-242-13113	Sequence 13113, A
21	94	4.0	2437	10 US-09-815-242-5834	Sequence 5834, Ap
22	93.5	4.0	429	10 US-09-815-242-13506	Sequence 13506, A
23	93	4.0	1295	10 US-09-726-949A-1	Sequence 1, Appl
24	92.5	4.0	432	10 US-09-919-172-9	Sequence 9, Appl
25	92.5	4.0	459	9 US-10-102-806-469	Sequence 469, Appl
26	92	4.0	3241	10 US-09-841-786-1	Sequence 1, Appl
27	91	3.9	1043	10 US-09-815-242-5194	Sequence 5194, Ap
28	91	3.9	1289	9 US-09-712-363-259	Sequence 259, App
29	91	3.9	2025	10 US-09-815-242-5703	Sequence 5703, Ap
30	90.5	3.9	550	9 US-10-125-692-22	Sequence 22, Appl
31	90.5	3.9	554	9 US-10-125-692-21	Sequence 21, Appl
32	90	3.9	375	9 US-09-738-628-6333	Sequence 6333, Ap
33	90	3.9	420	9 US-09-770-509-10	Sequence 10, Appl
34	89.5	3.9	483	8 US-08-834-666A-20	Sequence 20, Appl
35	89.5	3.9	628	10 US-09-881-736-4	Sequence 4, Appl
36	89.5	3.9	660	12 US-10-007-693-139	Sequence 139, App
37	89	3.8	99	10 US-09-925-302-445	Sequence 445, App
38	89	3.8	803	10 US-09-738-363-12	Sequence 12, Appl
39	88.5	3.8	4613	9 US-09-860-846-31	Sequence 31, Appl
40	88.5	3.8	4613	9 US-09-988-384B-31	Sequence 31, Appl
41	88.5	3.8	4613	9 US-09-836-821-31	Sequence 31, Appl
42	88.5	3.8	4613	10 US-09-861-289-31	Sequence 31, Appl
43	88.5	3.8	11877	9 US-09-860-846-6	Sequence 6, Appl
44	88.5	3.8	11877	9 US-09-836-821-6	Sequence 6, Appl
45	88.5	3.8	11877	10 US-09-861-289-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-5835
Sequence 5835, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5835
LENGTH: 2434
TYPE: prt
ORGANISM: Staphylococcus aureus
US-09-815-242-5835
Query Match 5.1%; Score 118.5; DB 10; Length 2434;
Best Local Similarity 21.2%; Pred. No. 0.72;

QY 184 LSI-----IAANVP-----YKFSNFIGLESRI-----SOGAA 212
 Db 890 MSAFEGCITPLVPSDEKRLKMGTRNDFTTYAGCGGLTPKQOYIADHTQVCANVM 949
 QY 213 TTSISDAKRAVDPI-----LLYCOLVMBETLLVDLAILYKRGMAEHV 255
 Db 950 YTALSRAKTRDRIHFVNTSANSASFWEKLDSTYLTFTLSVNEQALRE---YEPAAEAPI 1005
 QY 256 AS-----AVENANRVKELAADTL-----FLHKIIP 282
 Db 1006 REPEPOTMVCNEESVLEEKKEELLEKFREIHSSEHSGNCVOTEDTTLQLSHOAK 1065
 QY 283 EOALIGAVYH---PISASETS-----KALINVTYFGVDPVRPIGNRRYFT 327
 Db 1066 DETLMATIDARLKISNOETFRFELSKDIDGLVPLFYQAKMLP-----KERIPFS 1118
 QY 328 NSYWNYSICSEAYMNGVMFRGCSNVRNPNIIVSKMSDGFYTMENSDBRKLITYRKHDGW 387
 Db 1119 QEVWEA---CAHEVQSKYLSKRCNLINGTVRQS-----PDFENKIMVELKSO-- 1164
 QY 388 GNGTLEDDEPGQGHMRPILRHGKYW-----SKRRPNMFMWESSASGYIR 435
 Db 1165 -WYTKVVKRLG-----LPKIKPGQTIAAFYQOTVLMFGTARYMRPRAOFOPREYFI 1215
 QY 436 SWENNPGRPGHWSI 449
 Db 1216 NCETTPEDMSAMAL 1229

RESULT 12

US-09-134-001C-5282
 ; Sequence 5282, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lyna Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GPC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 5282
 ; LENGTH: 331
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-5282

Query Match 4.18; Score 94.5; DB 4; Length 331;
 Best Local Similarity 23.08; Pred. No. 0.28;

Matches 80; Conservative 50; Mismatches 111; Indels 107; Gaps 21;

QY 10 FIVLA-ITSAHGRSDVNSLITKVTALKEASGS-----NE-----AALEALE 52
 Db 13 FIMKVIYDEQTOQLDNEVILEVERKALQAFSEKKTITPLRYVLPEFQNRVLYVPLS 72
 QY 53-----GLKGEIOTKRPD--RVGAQTKILGSV-----GSALGKINSGDATKIISGCDI 97
 Db 73 DELNIYGLK-IVSFAPENSKKQKAT-ITGSVILSDYETGELSLIDGFLTKVRGA--- 127
 QY 98 VAGIATTFGGPVGMGIGAVASFVSSILSLF-TGSSAKNSVAAYVD-RALSXHRDEAIORH 155
 Db 128 ISGVATRY-----LAKENAKTLVIGAGVQAEGLIAILLAVRDIEK-----IH 170
 QY 156 AAGAKRPEASSAFIYVKKOOSN-----LTPDSLSIIR--ANVPYKFS----- 197
 Db 171 I--ASRTEPKAEKFAQNIIRNFNIKVSFBSADEAIDSDADIVYKATNANQVYTHSLHPG 228
 QY 198 ---NFIOLESRIISOGAATTSLSDAKRAVDLILYCOLVMBETLLVDLAILYKRG--- 250

Db 229 VHLNAGSFKRPMQEIPESTWLVANK-----IVSEMEALEET--GDLKIPQAGILTK 281
 QY 251 -----NAHVASAVENANRVKELAADL-----FLHKIIP 281
 Db 282 NMLHSELGDIISGERVRETEEVYKRSVGLAIVDIIIVAOYFYKLLI 329

RESULT 13

5183745-2
 ; Patent No. 5183745
 ; APPLICANT: DANCHIN, ANTOINE; GLASER, PHILIPPE; KRIN, EVELYN;
 ; BARZU, OCTAVIEN; LADANT, DANIEL; OULMAN, AGNES
 ; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
 ; BIOLOGICAL USES
 ; NUMBER OF SEQUENCES: 13
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/426,541
 ; FILING DATE: 25-OCT-1989
 ; SEQ ID NO: 2
 ; LENGTH: 1489
 5183745-2

Query Match 4.18; Score 94.5; DB 6; Length 1489;
 Best Local Similarity 25.08; Pred. No. 3.6;
 Matches 65; Conservative 49; Mismatches 115; Indels 31; Gaps 12;

QY 11 IVLAITSKHKRSDVNSLITKVTALKEASGSNEALEGLKGEI-----TKPRV 65
 Db 587 IALALAAAR-----GVTSGLVAGASAGAAALALALSPME-TYGLVQOSHVDOLDKL 640
 QY 66 GOATRIIGSVGSLGKINSGDATKIISGCLIVAGIATFGPGVMGIGA--VASFVSSI 123
 Db 641 AOBSSANGYEDDALLAOLYRKPT--AAECVAGVSAVSTVGAANSIAAASVYGAAPVAV 699
 QY 124 LSEFTGSANKNSVAVIDRALSKEH--RDEAIOHAAKGRDFAESSAFIYVKKOOSMLTD 181
 Db 700 TSLTIG--ALNGILRGVQOPIETKLANDYARKIDELGGQAYFEKN--LOARHEQLANSND 755
 QY 182 SDSLITIANVYVYKFSNFIGLESRIISG-----AATTSLSAKRAVD-FILLYCO-LVYM 235
 Db 756 GLRKLMDLQAGWNASVYIGVQTEIISKSALELAITGNADNLKSDVDFRFGGERVA 815
 QY 236 RETLLVDA-----ILYKRG 250
 Db 816 GQPVLDVYAAGIDIASRKG 835

RESULT 14

US-08-669-785-2
 ; Sequence 2, Application US/08669785
 ; Patent No. 6309648
 ; GENERAL INFORMATION:
 ; APPLICANT: Betsou, Fotini
 ; APPLICANT: Sebo, Peter
 ; APPLICANT: Guiso, Nicole
 ; TITLE OF INVENTION: Protective Epitopes Of Adenyl
 ; TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flunegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/669,785

QY 10 FIVLA-ITSAHGRSDVNSLITKVTALKEASGS-----NE-----AALEALE 52
 Db 13 FIMKVIYDEQTOQLDNEVILEVERKALQAFSEKKTITPLRYVLPEFQNRVLYVPLS 72
 QY 53-----GLKGEIOTKRPD--RVGAQTKILGSV-----GSALGKINSGDATKIISGCDI 97
 Db 73 DELNIYGLK-IVSFAPENSKKQKAT-ITGSVILSDYETGELSLIDGFLTKVRGA--- 127
 QY 98 VAGIATTFGGPVGMGIGAVASFVSSILSLF-TGSSAKNSVAAYVD-RALSXHRDEAIORH 155
 Db 128 ISGVATRY-----LAKENAKTLVIGAGVQAEGLIAILLAVRDIEK-----IH 170
 QY 156 AAGAKRPEASSAFIYVKKOOSN-----LTPDSLSIIR--ANVPYKFS----- 197
 Db 171 I--ASRTEPKAEKFAQNIIRNFNIKVSFBSADEAIDSDADIVYKATNANQVYTHSLHPG 228
 QY 198 ---NFIOLESRIISOGAATTSLSDAKRAVDLILYCOLVMBETLLVDLAILYKRG--- 250

APPLICANT: Turner, Nilgun E.
TITLE OF INVENTION: Plants Resistant to Infection by PVX
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/803,973
FILING DATE: 21-FEB-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/804,862
FILING DATE: 04-DEC-1991
APPLICATION NUMBER: US 07/771,912
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10541)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-803-973-2

Query Match 4.1%; Score 95; DB 1; Length 1456;
Best Local Similarity 19.0%; Pred. No. 3.1;
Matches 105; Conservative 67; Mismatches 188; Indels 194; Gaps 27;

QY 52 EGLKEIOTKPRVGOATKILGSVS-----ALGKINSG-DATKIISG---CLDIYA 99
Db 714 EYLTKITSLERKVAATCVIHGAGSGSKSHAIQKALREIGKSDITVVLPTNELRLDMSK 773
QY 100 GIATL-----FGG-----PYGMGIGAVASFYSSI-LSLFTGSSAK 133
Db 774 KVPNTEPYMKTEKALIGTGSIVIPDDYSKLPFGY-IEALICFYSKIKLIVLTGDSRQ 832
QY 134 NSVAVIDRALSKRDEAIO-----RHAAGA---KRDFAESSAFIOVMKQOSNLTDSD 183
Db 833 SVYHETADASIRHLGPATEYFSKYCRYLNTAHRNKKDLAN--MLGVYSERTGVEIS 889
QY 184 LSI-----IAANVP-----VYKFSNFIQLESRT-----SQGA 212
Db 890 MSAEFLBETPLVPSDEKRLKMYGTGRNDFTYAGCGILTRKQVIVLDHNTQVCSANVM 949
QY 213 TTSISDAKRAVDFT-----LYCOLVYMRETLLVDLAILYRKGNABHV 255
Db 950 YTALSRAIDRIHFVNTSANSASFMEKLDSTPYLKTFLSVADQALRE---YEPAEAPET 1005
QY 256 AS-----AVENANRVKELAAOTLD-----FLHKLIP 282
Db 1006 REPPOQHMCEVNESEVLEEKELKEFDREIHSESHGNSCVQTEDTTLQLSHQAK 1065
QY 283 EOALIGAVYH---PISASETS-----KAILNYTYFQVDPVPRIGNRKRYFT 327
Db 1066 DEFLMTATIDARKLISNOETFRFELSKKIDGDLVFLYQAKMGLP-----KERIPFS 1118
QY 328 NSYVNTYSICSEAVYMGVMEFGCSNVNPNIRVSKMSDGFYTMENSDBRKLYITRKDQGW 387

Db 1119 OEYWEA---CAHEYQSKYLSKSCNKLINGTVROS-----PDPEKINMFLSKO-- 1164
QY 388 GWTGLDEDDPDGDGMRFIP-LRHGKIYV-----SSKRPNNFYTMESSASGYR 435
Db 1165 -WYTKYEKLG-----LPRIKPGQRTIAAFYQOQVLMFGTWMRYMRWFOAFOPKEVEI 1215
QY 436 SWENNPQPGQHWST 449
Db 1216 NCETTPEDMSAMAL 1229

RESULT 11
US-08-803-972-2
Sequence 2, Application US/08803972
Patent No. 5792937
GENERAL INFORMATION:
APPLICANT: Braun, Carl J.
APPLICANT: Hemenway, Cynthia L.
APPLICANT: Turner, Nilgun E.
TITLE OF INVENTION: Plants Resistant to Infection by PVX
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/803,972
FILING DATE: 21-FEB-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/804,862
FILING DATE: 04-DEC-1991
APPLICATION NUMBER: US 07/771,912
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10541)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-803-972-2

Query Match 4.1%; Score 95; DB 1; Length 1456;
Best Local Similarity 19.0%; Pred. No. 3.1;
Matches 105; Conservative 67; Mismatches 188; Indels 194; Gaps 27;

QY 52 EGLKEIOTKPRVGOATKILGSVS-----ALGKINSG-DATKIISG---CLDIYA 99
Db 714 EYLTKITSLERKVAATCVIHGAGSGSKSHAIQKALREIGKSDITVVLPTNELRLDMSK 773
QY 100 GIATL-----FGG-----PYGMGIGAVASFYSSI-LSLFTGSSAK 133
Db 774 KVPNTEPYMKTEKALIGTGSIVIPDDYSKLPFGY-IEALICFYSKIKLIVLTGDSRQ 832
QY 134 NSVAVIDRALSKRDEAIO-----RHAAGA---KRDFAESSAFIOVMKQOSNLTDSD 183
Db 833 SVYHETADASIRHLGPATEYFSKYCRYLNTAHRNKKDLAN--MLGVYSERTGVEIS 889

APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hoffe, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
TITLE OF INVENTION: heteropolymers of compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
EARLIER FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 80
LENGTH: 3079
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-80

Query Match 4.3%; Score 100.5; DB 4; Length 3079;
Best Local Similarity 22.5%; Pred. No. 2.9; Mismatches 126; Indels 113; Gaps 18;
Matches 83; Conservative 47;

QY 38 KESGSGEALLEGKGTQKPRD-----VCOATKILGSGV 76
DB 55 REAGGAEYWRKALDGTATDLPDRARADAGARGRAHATLPKLTGALRLAERG 114
QY 77 SALGKLNSGDATKII-----SCLDIYAGI-----ATTFGG-----PVGNGIGAVAS 118
DB 115 TTFPSVLSALTYLHRASQSDLVGVGSPAGRHDESSARAFGYVOMLTVRVALLRGAAS 174
QY 119 FVSSILSLEFGSSAKNSVAVIDRALS--KRDDEATORHAGAKRDEASSAPITQVKKQ 176
DB 175 FDA-----LVARVDAFLDALAHGDSALRHLLAROGEGORALDVAFAFO 221
QY 177 SN--LTFSDLSITIANVPYKFSNFTGOLESRTSG--AATSLSDAKRAVDF--ILLYCO 231
DB 222 STPPSLDALSLALA-----IGVGDVRIAGOELELTVIADQAAAEFDLALFAAE 270
QY 232 L---VVMR---ETLVLDLAILYRKNAER-----VASAVENAMVNNELADTDLFLKLI 281
DB 271 LDAGIALREYDQGLDPATIERM--ARHFVULLESVAHEPGRPSELNM--ISDAERALL 327
QY 282 -----PEGALIGAVYHPIASSETSRAILNTYKYEFGVPDVPPIGNRRYK 325
DB 328 LDDWSGAAARQASAPAPACVHALFEAHARQPDATALEF-----GHQRT 374
QY 326 FTN-SYWN 333
DB 375 YAGLSTWST 383

RESULT 7
US-09-255-829-18
Sequence 18, Application US/09255829
Patent No. 6461617
GENERAL INFORMATION:
APPLICANT: Shone, Clifford Charles
APPLICANT: Quinn, Conrad Padraig
APPLICANT: Foster, Keith Alan
TITLE OF INVENTION: Recombinant Toxin Fragments
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/255,829
FILING DATE: 23-FEB-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB97/02273
FILING DATE: 22-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/782,893
FILING DATE: 27-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 1581.0130002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-255-829-18

Query Match 4.2%; Score 97.5; DB 4; Length 1013;
Best Local Similarity 20.9%; Pred. No. 0.91; Mismatches 154; Indels 163; Gaps 25;
Matches 98; Conservative 55;

QY 9 LFIYLAITSKRGCK-----RSDVNSLL-----TKVET-----ALKEAGSGNEAA--LEALRG 53
DB 557 MEHYLAOEFEHKSRIATLNSVNEALLNPSRYTFSSDYKVKVKAETAEAMFGLGWQD 616
QY 54 L-----KGEIOTKPDVGOATKILGSGVSGALGKNSGD-----ATKIIISGC---L 95
DB 617 LYVDFDESEVST--TDKADITITIIPIYIGPA--LNIQNMLYKDDFVAGALJFSGAVILL 672
QY 96 DIVAGIATTEGGPVGNGIGAVASFWSSILSFTGSSAKNSVAAY--IDRALSKHRD----- 149
DB 673 EFIPETAI-----PV--LGFPAIVSYIA-----NKVLTVQITIDALSKRNEKME 715
QY 150 -----EAIORHAGAKRDE--AESSAFIOVMKQOOSNLTSD 183
DB 716 VKIYLVNMLAKVNTQIDLIRKKMEALENOEATKAITIYQNOYTEEEKNNINFNID 775
QY 184 L-----SIIANVPYKFSNFI--G--QLESRIISGGAATSLSDAKRAVDFTILYCOLV 234
DB 776 LSKKNESINKAMINIKFLNOCSTYILNSKIPYG-----VKRLDF----- 818
QY 235 MRETLVDLAILYRKNAEHAASAVEN--ANRYNKEADTDLFLKLIPEOALIGAVYHP 293
DB 819 --DASLKDALKKIYINDNGTLIGVDRLKDKVNNNTLSDIPOLSKYVDNRLLSTFEY 876
QY 294 ISASETSKAILNTYKYEFGVPDVPPIGNRRYKFTNSYWN--YSICSEAYMGNYMFGCSN 352
DB 877 IKSGINSPGAHYAQHDEAVD-----NKFKEQOONAFYEI----- 911
QY 353 VRNPNIRSKMSDGYFMENSDRKLITKHDCQWGTGIDEDGDGGM 402
DB 912 LHLPLN-----NEQDNNAFIQ-----SLKDDPSOGANL 939

RESULT 8
US-08-968-685A-10
Sequence 10, Application us/08968685A
Patent No. 6214981
GENERAL INFORMATION:
APPLICANT: TUCKER, KENNETH
APPLICANT: PLOSILA, LAURA


```
FILE REFERENCE: 210121.469C7
CURRENT APPLICATION NUMBER: US/09/620.412C
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 191
LENGTH: 977
TYPE: PRT
ORGANISM: Chlamydia
US-09-620-412C-191

Query Match
Best Local Similarity 21.3%, Pred. No. 0.024; DB 4; Length 977;
Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 26;

21 GKRSDV-----NSLTKVETALKEASGNEALELEGLKEIOTKPRDVGATKILGSYG 76
267 GNRGNIVYNNRCFNVEFASSEAS-----DGAIKVTTRLDVTGNRGRIFFS-- 314
77 SALGRLNGDGTKIISGCLDIYAGIATFGGPGVGMGICGAVASFVSSILSLFGSSAKNV 136
315 -----DNITKNYGAI---YAPVYTLVDNGPPTFTNNIANNKG 349
137 AAVIDRALSKHDEAIORHA-----AGAKR---DFAESSAFIOY 172
350 GAIYIDGTSNKRISA-DKHAIFENNTYNTNANGTSSANPPRNMATTVASSSGEILL 408
173 MKQGS-NLTDSD-LSIIANPVYKFSNFTGLESRIISOGAFTSLSDAKRAVD----- 224
409 GAGSSQNIIEFYDPIEVSNAGV-SFNKEADQTSVFEFGAVNSADEHQRLQRTKTPAP 467
225 -----FILL--YCOLVYMRRL--LVDL---ALL--YRKNAEIVASAVENARKVKE 268
468 LTLNGLFICIEDHAQLTVNRFTQGTGVVSLGNGAVLSCTKNKTGD---SASNASITLKH 523
269 LAADLDFLHKLIPQALIGAVYHPISASETSKAILNT---KTFGVDPVPRPI-----G 320
524 IGLN-----LSSILKSGAEI-----PLMVEPTNNSNNTADPAAFTSLSDVLSLIDYG 574
321 NRRYKFTNSYVNTYSIGSEAYGNVYMFRCSSVNRPNIRVSKMSDGFYTWENSRRKRLYI 380
575 NSPYESTDL--THALSSQ-----PMLSISEADNQLQSENIDFSGLVN 615
381 TKHD-QG---WGNG--TLDEDDPGDQGHMR-----PIPL-RHCKYV 414
616 PHYGQGLMTWGMARTQDEPASSATITDPOKANRPHRLTLTWLPAGVVPSPKRSPLI 675
415 SSKRPNFMFMWESSASGYIRSMENNPQGGHMSIT 450
676 ANTLINGNMILATESLKN-----SALTPSGHPFWGIT 707

RESULT 5
US-07-689-008-2
Sequence 2, Application US/07689008
Patent No. 5268274
GENERAL INFORMATION:
APPLICANT: Ben-Bassat, Arie
APPLICANT: Calhoon, Roger D
APPLICANT: Fear, Anna L
APPLICANT: Gelfand, David H
APPLICANT: Meade, James H
APPLICANT: Tal, Rony
APPLICANT: Wong, Hing
APPLICANT: Ben-Ziman, Moshe
TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Enersen
STREET: Three Embarcadero Center
CITY: San Francisco
STATE: California
```

```
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/689, 008
FILING DATE: 19910422
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 337,194
FILING DATE: 12-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 496,236
FILING DATE: 23-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Lisabeth Felix
REGISTRATION NUMBER: 31547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 393-2000
TELEFAX: (415) 393-2286
TELEX: 340817 MACPAG SFO
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3031 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-689-008-2

Query Match
Best Local Similarity 4.7%, Score 110; DB 1; Length 3031;
Matches 79; Conservative 43; Mismatches 127; Indels 112; Gaps 14;

24 SDVNSLTKVETALKEAS---GSNEALEAL-----EGLKEIOTKPRDVGATKILGSYG 76
1423 TVESTLIGGVAAIMLEASPLASGRVYALLSGDQGLNNVLQILAQRNQA-KIQGDVY 1481
77 SALGRLNGDGTKIISGCLDIYAGIATFGGPGVGMGIC-----AVASFVSSIL 124
1482 LAHG-----DLTSYSSPLYTV-----GTPMLLEPDWYMHNPFRVIVYGLGCLT 1529
125 SLFTSSSAKNSVAAYVIDRALSKHDEAIORHAAGKRDPAESAFIQ---YMKQGSNLT 180
1530 -----IYAAVVRRLAHN---ALNR-----RRELQERORTMNRVYLSLGA 1570
181 DSDLSIIANPVYKFSNFTGLESRIISOGAFTSLSDAKRAVDFTLLYCOLVYMRRL 240
1571 ASSCHTVLYAVYVAR-----AQOASTAMTTATATATA-----PQILL 1609
241 VDLAILYRKGAHEVAVASAVENANRV-----NKELAADLDFLHKLI 281
1610 QOAREFLQOOQYDANARQALONARLAPNSPDYLEVIGEYQTAIGRREAAADLRHLOQVA 1669
282 PQOALIGAVYHPIS-----ASETSKAILNTYKTFGVDPVPRPIGNRRY 324
1670 PGSAAGNLDLILSERAIQSPLSOTIRSLAGSGQNNQAVAGYQKLFHGKPPHSLAVEY 1729
325 K 325
1730 Q 1730

RESULT 6
US-09-413-814-80
Sequence 80, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
```

```

QY 21 GKRSDV-----NSLITVEFALKEAGSGNSEALALEGLKGEIOTKPDVVGQATKILGSGV 76
Db 254 GNRGINIVFYNNRCFKNEVETASSEAS-----DGAIKVTRRLDVTGNRGRIFFS-- 301
QY 77 SALGRINSGDATKIISGCDIYAGIATFEGPVGMGICAVASFPVSSILSTFGSAAKSV 136
Db 302 -----DNITKNTGAI-----YAPVTVLVDNCPTEYFINNIANNKG 336
QY 137 AAVIDRALSKHDEAIORRA-----AGAKR-----DEAESATQY 172
Db 337 GAIYIDGTNSKRISA-DRAIIIFENIYITVNTNANGSTSNAPPRRAIIVASSGEBIL 395
QY 173 MGOOS-NLTDSP-LSIIAANPVYKFSNFQILESRISOGAATYLSLDARAVD----- 224
Db 396 GAGSSONIFEDPIEVSNAGVSV-SFNKEADQGTGSVFSGATVNSADFHORNIQTTPAD 454
QY 225 -----FILL--YCOOLVYMRTEL--LYDL-----AIL--YRKGNAEHVASAVENARVKE 268
Db 455 LTLNSGFLCIEDHCOLVYNRFTQGTGGVAVSLGNGVILSCYKNGTGD-----SASNAJSLIHK 510
QY 269 LAADTDLFLHKLIPEDQALIGAVYHPISASESKAILNYT-----KYRGVDPYPPPT---G 320
Db 511 IGLN-----LSSILASGAEI-----PLWVEPTNNSNNYTDIATATSLDYSVKLSLDDYG 561
QY 321 NRRKFTINSYNTYISCSSEAYMGNYMERGCSNVRNPNIIRVSKMSDGFYTMENSDDRKLKI 380
Db 562 NSPESRDUL--THALSSQ-----PMLSTSEASDNQLOSENIDFSGLNV 602
QY 381 TKHD-OG--MGWG-TLDEDPGOGGHR-----FIPL-RHQKYAV 414
Db 603 PHYMOGLMTMGWAKYTDDEPPASATITPOKANREFHRLILLTLWLPAGVVPSPKRRSPLI 662
QY 415 SSKRMPWMEFYMMESSASAGYIRSWENNPDGQGHSTI 450
Db 663 ANLTMGNMLATSLSKN-----SHELTPSGHPFWGIT 694

```

```

0Y 21 GKSSDV---NSLLTVEVRLAEASSNMAALEEGLEKGEIOTKPDREVGATITLGSVG 76
Db 267 GNGNIVTFYNNRCKFVNEVJASBASS-----DGAIKVTRILDTVGNRGHIFFS-- 314
0Y 77 SALGLNSGDATKIISGCLDIYAGIATFEGGVGIGAVASFVSSILSTFGSARNSV 136
Db 315 -----DNITKNVGGAI---YAPVVTIYLDNGETYIINNIAKNG 349
0Y 137 AAVIDRALSKHDEALIORHA-----AGAKR---DEASSAFIOY 177
Db 350 GAIYIDGTSNKSISA-DRIALIFENIYTVTNMGSTSNPNPRRAIYVASSSGEILL 408
0Y 173 MKOQS-NLTDSD-LSIIAANPVYKFSNFIGOLESRISOGAATYLSDAKRAYD----- 224
Db 409 GAGSSONLIFYDPIVSNAGVSV-SFNKEADQTSVAFSGATVNSADFHQRNLOTKTPAP 467
0Y 225 -----FILL---YCOLVYMRRL---LYDL---AIL---YKRGNAEHVASAEVANRYNKE 268
Db 468 LFLNSGFLCIEHQAOLITVNRFTQGTGGVAVSLGANGVLSCYRKNGTGD---SASASITLKH 523
0Y 265 IAAOTLDPLPHKILPEBQALIGAVYHPISASETSUKALNYT---KYFCVDPVPRPI----G 320
Db 524 IGLN---LSSILKSGAEI---PLWVEPTNSNNYTMADPAITSLSDVYKSLDIDYG 574
0Y 321 NRRYKFTSNVYNTSISESEAVMGVYMRGCSNYPANPIRYSKSDCFYTWENSDDRKLXI 380
Db 575 NSPYSTSLD---TALSSQ-----PWLISEASDQLOSENIDEFGLNAV 615
0Y 381 TKHD-QG---WGWG-TLDEPDGDOGHR-----RPL-RHCKVYV 414
Db 616 PHYMOGILMTGAKTKODPEPASATITDPQKANREHRLULLTWLPAGIYVSPKHSRPI 675
0Y 415 SSKRMPNWFYMESSASGAYIRSMENPNPGOGHHST 450
Db 676 ANTIAMGNMLATESLKN---SALTPSPGHPFGIT 707

```

RESULT 4
US-09-620-412C-191
; Sequence 191, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Flinn
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

Gencore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 11:14:52 ; Search time 30 Seconds
(without alignments)
441.344 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322
Sequence: 1 MLTKLPLWLFVLAITSAKH.....SGYISWENNPGQGHWSIT 450

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*
5: /cgn2_6/prodata/1/1aa/PCRTUS_COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112.5	4.8	964	US-09-556-877-177	Sequence 177, App
2	112.5	4.8	964	US-09-620-412C-177	Sequence 177, App
3	112.5	4.8	977	US-09-556-877-191	Sequence 191, App
4	112.5	4.8	977	US-09-620-412C-191	Sequence 191, App
5	110	4.7	3031	US-07-689-008-2	Sequence 2, App1
6	100.5	4.3	3079	US-09-413-814-80	Sequence 80, App1
7	97.5	4.2	1013	US-09-255-829-18	Sequence 18, App1
8	97	4.2	2123	US-08-968-685A-10	Sequence 10, App1
9	95	4.1	1434	US-08-989-925-1	Sequence 1, App1
10	95	4.1	1456	US-08-803-973-2	Sequence 2, App1
11	95	4.1	1456	US-08-803-972-2	Sequence 2, App1
12	94.5	4.1	1456	US-08-803-972-2	Sequence 2, App1
13	94.5	4.1	1489	US-09-134-001C-5282	Sequence 5282, App
14	94.5	4.1	1706	US-08-669-785-2	Sequence 2, App1
15	94.5	4.1	1794	US-08-669-785-2	Sequence 2, App1
16	93.5	4.0	580	US-08-591-079-2	Sequence 2, App1
17	93	4.0	1296	US-08-460-604A-28	Sequence 28, App1
18	93	4.0	1296	US-08-405-96A-28	Sequence 28, App1
19	93	4.0	1296	US-08-915-136-28	Sequence 28, App1
20	92.5	4.0	432	US-08-705-660-18	Sequence 18, App1
21	92.5	4.0	432	US-08-989-045-18	Sequence 18, App1
22	92.5	4.0	689	US-08-684-865-16	Sequence 16, App1
23	92.5	4.0	689	US-08-124-491-16	Sequence 16, App1
24	92.5	4.0	907	US-09-255-829-16	Sequence 16, App1
25	92	4.0	497	US-08-740-223A-14	Sequence 14, App1
26	92	4.0	497	US-09-709-188-14	Sequence 14, App1
27	92	4.0	593	US-08-591-079-8	Sequence 8, App1

28	92	4.0	593	2	US-08-591-079-10	Sequence 10, App1
29	91.5	3.9	491	4	US-08-740-223A-13	Sequence 13, App1
30	91.5	3.9	491	4	US-09-709-188-13	Sequence 13, App1
31	91.5	3.9	937	1	US-08-253-155A-31	Sequence 31, App1
32	91	3.9	585	2	US-08-426-125-2	Sequence 2, App1
33	91	3.9	585	2	US-08-455-355-2	Sequence 2, App1
34	91	3.9	1026	2	US-08-542-003-6	Sequence 6, App1
35	91	3.9	1026	2	US-08-322-760A-6	Sequence 6, App1
36	91	3.9	1026	4	US-09-236-949-6	Sequence 6, App1
37	90.5	3.9	1705	4	US-08-669-785-4	Sequence 4, App1
38	90.5	3.9	10182	4	US-09-134-001C-3159	Sequence 3159, App
39	90	3.9	619	1	US-07-762-132A-2	Sequence 2, App1
40	90	3.9	619	1	US-08-301-722A-4	Sequence 4, App1
41	89.5	3.9	535	4	US-09-117-860-18	Sequence 18, App1
42	89.5	3.9	15281	2	US-08-471-119A-2	Sequence 2, App1
43	89	3.8	401	2	US-08-591-079-4	Sequence 4, App1
44	89	3.8	803	1	US-08-158-232-10	Sequence 10, App1
45	89	3.8	803	1	US-08-304-626-10	Sequence 10, App1

ALIGNMENTS

RESULT 1	US-09-556-877-177	Sequence 177, Application US/0955687
Patent No. 6432916		
GENERAL INFORMATION:		
APPLICANT: Probst, Peter		
APPLICANT: Bhatia, Ajay		
APPLICANT: Skelky, Yasir		
APPLICANT: Flind, Steve		
APPLICANT: Maisonneuve, Jeff		
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND		
FILE REFERENCE: 210121.469C5		
CURRENT APPLICATION NUMBER: US/09/556,877		
CURRENT FILING DATE: 2000-04-19		
NUMBER OF SEQ ID NOS: 305		
SOFTWARE: FastSeq for Windows Version 3.0/4.0		
SEQ ID NO 177		
LENGTH: 964		
TYPE: PRT		
ORGANISM: Chlamydia		
US-09-556-877-177		
Query Match	4.8%; Score 112.5; DB 4; Length 964;	
Best Local Similarity	21.3%; Pred. No. 0.023;	
Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;		
QY 21 GKRDV-----NSLTKVETALKEAGSGNEALELDELKGEIQKPPRGQATKILGSVG 76		
DB 254 GNRNIEFYNNRCKPKNETASSSENS-----DGAIRKVTTRLDVGTGRGIFPS-- 301		
QY 77 SALGKLSGDAKTKIISGCLDIAGIATFGGPGMGIGAVAFSSILSFTGSSAKNSV 136		
DB 302 -----DNTKNYGCAT---YAPVTVLVNCPYTFINNINANKG 336		
QY 137 AAVIDRALSKRDAIQRHA-----AGAKR-----DFAESSAFIOV 172		
DB 337 GAIYIDTGSNSKISA-DRHAIFENIIVTNTANGTSTSNPFRRAITVAASSGCEILL 395		
QY 173 MKQOS-NUTBSD-LSIITANVPYKFSNFIGLESRSOGAATLSLDARAVD----- 224		
DB 396 GAGSSQNLIFDPIEVSNAGVS-SEFKKADQTSVVFSGATVNSADPHQRNLQTKTPAP 454		
QY 225 -----FLL--YCOLVMEETL---LVDL---AIL--YRKGNAEHVAVENANRYNKE 268		
DB 455 LTLNGLICIEDHQVLVNRRTQGVVSLGNGAVLSCYKNGTID-----SASNASITLKH 510		
QY 269 LAADTLPLHLKILPEQALIGAVYHPIASSETSKAILNT-----KYGVDPVPRPI---G 320		
DB 511 IGLN-----LSSILKSGAEI-----PLWVEPTNNSNMYTDTATATFSISLIDYKSLIDY 561		

Db 492 ICSDKTGT-----LTLNOMTVV-ET-----YAGSKMDVA---DNPSGLHPKLV A--- 532
Oy 275 DELHKLPE---QALIGAVHP-----ISASETSKALINTKTYFGV 312
Db 533 -----LISEGVAQNTTGNIFHPKDGGEVEISGSPTEKAILSWAYKLG M 575

Search completed: May 19, 2003, 11:16:24
Job time : 87 secs

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.78; Score 110; DB 21; Length 1086;
Best Local Similarity 22.88; Pred. No. 2.4;
Matches 93; Conservative 51; Mismatches 118; Indels 146; Gaps 20;

QY 9 LFIYLAITSAKHGRSD-----VNSLTKVEPLKBPASGSENALEBGLK 55
DB 210 ILIITAATSLALGKTEGKAKGMDGGSIAFAVLIVYTVASPYRSLO--PQNLDER 267
QY 56 GEIQFKPRDVGQATKI-----LGSVSGALGKINSQDATKI-----90
DB 268 RNIOLEVMRGRVTAKISIDYVVGVDVPIPLRIGDQVPADGVLSHSLAIDESSMTGESKI 327
QY 91 -----ISGC-----LDIYAGIATFG-----GPGMGIGAVAS 118
DB 328 VHKDOKSPLMSGCKVADGVGNMLVGTGVIWTEGMLMASTSEDTGETPQVRLNGIAT 387
QY 119 FVSSILSLFTGSSAKNSAAVAIDRLSKHRDEAIORAAAGAKRPAFASAFIQVKKOOSN 178
DB 388 FI-GIVGL-----SVALVVLVAL-----LVRYFTGTOTPDNATQIKXTTISD 431
QY 179 LTDS-----DLSITANP-----YKFSNFIGO--LESRIS---OGAATY 214
DB 432 IVDCKVKIETIAVIVVAAVEGLPLAVTTLTAYSMKRMADKALVRLSACETMGSAVY 491
QY 215 SLSDAKRAVDFTILYCOLVAVRETLVLDAIILYKRNAEHVASAVERNANRYNKEIADTL 274

PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159293.
 PR 14-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159337.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 28-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 29-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match
 Best Local Similarity 4.7%; Score 110; DB 21; Length 1073;
 Matches 93; Conservative 51; Mismatches 118; Indels 146; Gaps 20;

DB 9 LEIVLAITSAGHGRSD-----VNSLTKVETALKEASGSENAALALBGLK 55
 DB 197 ILITIAVTSLALGKTGTEGKGLDGSIAFAVLTVVAVSDYRSLQ--FQNLNDEK 254
 QY 56 GEIOTPRDRGQATKI-----LGSVGSALGKLGNDGDKTKI----- 90
 DB 255 RNIOLEVMRGGRTVKISIDVYVGVPIPLRIGDQVPADVLLGSLALDESSMGESKI 314
 QY 91 -----ISGC-----LDIVAGIATFG-----CPVGMGIGAVAS 118
 DB 315 VHKDQSPFLMSGCKVADGVGNMLVTGVTGINTENGILMASISEDGETPPLQVRLNGLAT 374
 QY 119 FVSITSLFTGSSAKNSVAVIDRALSKHRDEAIORHAAGAKKRDPASSAFIOVMKQSN 178
 DB 375 FI-GIYGL-----SVALLVVAL-----LVRIETGTDTDNGATQFIKGTISID 418
 QY 179 LKDS-----DISTAANP-----VYKFSNFIGQ--LESRS--OQAAT 214
 DB 419 IYDDCVKIFTTIAVTIVVAVPGLPLAVTLTLAYSMRKMAADKALVRLSACETMGSAT 478
 QY 215 SLSDAKRAVDLILYCOLVYMETLLVDLALYKRGNAEHVASAVENANRYNKELEADTL 274
 DB 479 ICSDKTGT-----LTLNMTVY-ET-----YAGGSKMDVA--DNPSGLHPIVA--- 519
 QY 275 DFLKLIP-----QALIGAVHP-----ISASETSKAILIYTKYFGV 312
 DB 520 -----LISEGVAQNTIGNIFHPKDGGEVIESSPTEKAILISMAVYKIGM 562

RESULT 15
 AAG49830 standard; Protein: 1086 AA.
 AAG49830;
 18-OCT-2000 (first entry)
 Arabidopsis thaliana protein fragment SEQ ID NO: 63079.

XX XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 PN EPI033405-A2.
 PD 06-SEP-2000.
 XX XX
 PF 25-FEB-2000; 2000EP-0301439.
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 16-APR-1999; 99US-0128714.
 PR 19-APR-1999; 99US-0128845.
 PR 21-APR-1999; 99US-0130077.
 PR 23-APR-1999; 99US-0130449.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 07-MAY-1999; 99US-0132486.
 PR 11-MAY-1999; 99US-0132487.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 25-MAY-1999; 99US-0135629.
 PR 27-MAY-1999; 99US-0136021.
 PR 28-MAY-1999; 99US-0136392.
 PR 01-JUN-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137522.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139492.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148541.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151330.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.

[illegible]

XX DE Protein encoded by Chlamydia trachomatis pmpe gene.
 XX KM Chlamydia; vaccine; infection; fusion protein; antigen;
 KM pelvix inflammatory disease; trachoma; atherosclerosis; heart disease;
 KM acute respiratory tract infection; Cap1; CT529; OMC8;
 KM polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
 XX OS Chlamydia trachomatis.
 XX PN MO200140474-A2.
 XX PD 07-JUN-2001.
 XX PF 04-DEC-2000; 2000MO-US32919.
 XX PR 03-DEC-1999; 99US-0454684.
 XX PR 19-APR-2000; 2000US-0556877.
 XX PR 20-JUN-2000; 2000US-0598419.
 XX PA (CORI-) CORIXA CORP.
 XX PI Probst P, Bhatia A, Skelky YAM, Fling SP, Scholler J;
 XX DR WPI; 2001-374831/39.
 XX PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
 PT inflammatory disease, trachoma, acute respiratory tract infections,
 XX atherosclerosis and heart disease -
 PS Claim 2; Page 216-218; 295pp; English.
 CC The present sequence is provided in a specification relating to
 CC compounds and methods for the treatment and diagnosis of chlamydial
 CC infection. The compounds provided include polypeptides and fusion
 CC proteins comprising immunogenic portions of Chlamydia antigens
 CC and DNA sequences encoding such polypeptides. They are useful for
 CC vaccinating against chlamydial infection, which causes pelvic
 CC inflammatory disease, trachoma, acute respiratory tract infections,
 CC atherosclerosis and heart disease.
 XX CC
 XX CC
 SQ Sequence 977 AA;
 Query Match 4.8%; Score 112.5; DB 22; Length 977;
 Best Local Similarity 21.3%; Pred. No. 1.2;
 Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

DB 575 NSPESTDLD---THALSSQ-----PMLTSEASDNOQSENIIDFSLN 615
 QY 381 TKHD-OG---WGWG-TLDEPDGDGHR-----FIL-RHGKRV 414
 DB 616 PHIGWGLMTWGMWAKTODPEASSATTTDPQKRNFRHTLLTLWLPAGVSPKRRSPLI 675
 QY 415 SSKRPMWFMWESSAGYIRSMENPNPQGWGKIT 450
 DB 676 ANTLGMNMLATESLKN-----SALTPSGHPFWGIT 707
 RESULT 12
 ID ABB94179
 XX ABB94179 standard; Protein; 977 AA.
 XX AC ABB94179;
 XX DT 05-JUN-2002 (first entry)
 XX DE Chlamydia protein sequence SEQ ID NO:191.
 XX KM Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
 KM antigen; antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.
 XX OS Chlamydia sp.
 XX PI Fling SP, Skelky YAM, Probst P, Bhatia A;
 XX DR WPI; 2002-179901/23.
 XX PT Novel compositions comprising Chlamydia Cap1 protein and its use in the
 PT treatment of Chlamydia infection -
 XX XX
 PS Example 1; Page 247-249; 537pp; English.
 CC The present invention describes compositions comprising a Chlamydia Cap1
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.
 CC Chlamydia DNA and protein sequences from the present invention can have
 CC antibacterial and immunostimulant activities, and can be used in
 CC vaccines. Compounds from the present invention can be used for eliciting
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell
 CC response or inhibiting the development of a Chlamydia infection in an
 CC animal. Methods from the present invention can be used for detecting the
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia
 CC infection. AB192394 to AB192709 and ABB94096 to ABB94374 represent
 CC sequences used in the exemplification of the present invention.
 XX CC
 XX CC
 SQ Sequence 977 AA;
 Query Match 4.8%; Score 112.5; DB 23; Length 977;
 Best Local Similarity 21.3%; Pred. No. 1.2;
 Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

Query Match 4.8%; Score 112.5; DB 23; Length 964;
Best Local Similarity 21.3%; Pred. No. 1,2; Indels 161; Gaps 28;
Matches 110; Conservative 61; Mismatches 184;

OY 21 GKRSVD---NSLTKRVEALKEAGSGSNDALEALEGLKGEIOTKRDVGRQATKIIGSVG 76
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 254 GKRGINVFNNCKFKNVETASBEAS-----DGAIKVTITRLDVGNRGRTFS-- 301
OY 77 SALGKLNSGDATKKIIISGICDIIVAGIATTFEGPVGMGIGAVAFVSLSILFTGSSAKNSV 136
DB 302 -----DNITKKNNGAI--YAIVTVLVNDGPPIFYFINNIANNKG 336
OY 137 AAVIDRALSKHDEAIORIA-----AQAKR---DEPAESSAFIOV 172
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 337 GAIVYIDGTNSNRISA-DRAHAIFFENETIVNTNANGTSTANPPRRNAITVAASSGEILL 395
OY 173 MKQOS-NLTDSO-LSTIANNPVKFNSNTIGOLESRISOGANTTSLSDAKRAYD----- 224
DB 396 GAGSSONLIETYPTEIVSNSGVSV-SFNKEADQTGSVFSGAVVNSADFQRNLQTTTPAP 454
OY 225 -----FIL--YCQLVVRRETL--LVDL---AIP--YRKGNAEHVASAVENANKVKE 268
DB 455 IFLNSGFELCIEHDQAOLTVRFOTGGVSVLSLGAVAYLCSTCKNGTD----SASNASTLLKH 510
OY 269 LAADTLDLFLHLKLIPEQALIGAVYHPDISASETSKAILNYT---KFGVPDYPRPI---G 320
DB 511 ICLN-----LSSTLKSGAEI-----PLWMEPPNNSNNYTADTAAPFTSLSDVKTSLDDYG 561
OY 321 NRRYFTNSYNWTYSICSAIYMGNTWFRCGCSNVRPNNINIVSKMSDCGYTBMBNSDRKLYI 380
DB 562 NSPYESTDL--THAISQG-----PMLSISEASDMQLOSENIDEPSGLMV 602
OY 381 TKHD-OG---WGMG-TLDEDPGDGHR-----PIPL-RHGKYAV 414
DB 603 PRYGMQGLTWKMGATQDEPPASSATIPTPOKANRHFRLLLTWLPAGIVPSPKRNKSPLI 662
OY 415 SSKRMPNMWEMESSASGYIRSMENPNPDGHMSIT 450
DB 663 ANTLGNNMLATESLKN---SAELTPSGHPMWGIT 694

RESULT 10
AABI3640
ID AABI3640 standard; Protein; 977 AA.
AC AABI3640;
DT 02-FEB-2001 (first entry)
C. trachomatis pmpe gene amino terminus minus signal sequence protein.

KM Chlamydia infection; sexually transmitted disease;
KM pelvic inflammatory disease; PID; tubal obstruction; infertility;
KM trachoma; blindness; acute respiratory tract infection;
KM atherosclerosis; coronary heart disease; antibacterial.
OS Chlamydia trachomatis.
XX WO200034483-A2.
XX PD 15-JUN-2000.
XX PF 08-DEC-1999; 99WO-US29012.
XX PR 08-DEC-1998; 98US-0208277.
XX PR 08-APR-1999; 99US-0288594.
XX PR 01-OCT-1999; 99US-0410568.
XX PR 22-OCT-1999; 99US-0426571.
XX (CORI-) CORIXA CORP.
PI Probst P, Bhatia A, Skelky YAW, Fling SP, Jen S, Stromberg EJ;
XX

DR WP1: 2000-431303/37.

XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection

PT Comprises immunogenic portion of Chlamydia antigen, which comprises

PT amino acid sequence encoded by polynucleotide sequence -

XX

P5 Claim 2; Pages 210-212; 256pp; English.

XX

CC The present invention relates to new nucleic acid sequences and the

CC proteins encoded by the nucleic acid sequences. The encoded proteins

CC comprise an immunogenic portion of a Chlamydia antigen. The encoded

CC proteins are useful for the serodiagnosis and treatment of Chlamydia

CC infection. Chlamydiae are intracellular bacterial pathogens that are

CC responsible for a wide variety of human infections. C. trachomatis

CC infection is one of the most common sexually transmitted diseases and can

CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction

CC and infertility. Trachoma due to ocular infection with C. trachomatis is

CC the leading cause of preventable blindness worldwide. C. pneumonia is a

CC major cause of acute respiratory tract infections in humans and is also

CC thought to play a role in the pathogenesis of atherosclerosis and

CC coronary heart disease. The present sequence is a protein isolated in the

CC present invention.

CC

XX Sequence 977 AA;

SQ

Query Match 4.8%; Score 112.5; DB 21; Length 977;
Best Local Similarity 21.3%; Pred. NO. 1.2;
Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

OY 21 GKRSVY---NSLTKVETALKEASGSNEALEGLEKGIQTQPDRYGATKLISVG 76
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 267 GNRGNIVFYNNCFNVEVASSEAS-----DGAIKYTTLDTVGNRGRIFFS-- 314
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 77 SALGLNLNGSDAKRKIIISGLDIYAGIATFPGGVGIGAVASFVSIILSLFGSSAKXSV 136
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 315 -----DNITKNYGAI---YAPVVLTVNGPFTYFNINANNNG 349
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 137 AAVIDRALSKHDEAIOHA-----AGAKR---DFEASSAFIOY 172
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 350 GAIVIDGTSNKSISA-DHAILFENENITYNTNANGSTTSANPPRNATTVASSGEILL 408
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 173 MKQOS-NLTDS-LIITAANPVYFSNFIGOLESRIISOGAATYSLSDAKRAVD----- 224
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 409 GAGSSQNLIIFYDPREVSNAGVSV-SFNKEADQTGSVVFSGAVVNSADEHQRMLOKTTPAP 467
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 225 -----FIL--YCOLVMRETL---LVDL---ALL--YKGMNAHVSAVENAKRVNKE 268
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 468 ILSNGFLCIEDHAOLTYNRFOTGGVYSLNGAVLSCCKNKGTG---SASNASTYLKH 523
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 269 LAADTLDLFHLKLIPQALIGAVYHPISASETSKAILNTY---KYGYEDVVRPI---G 320
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 524 IGLN----LSSLTKSGAEI-----PLWVEPFINNSNNTTADTAATFSLSDVYLSLIDYG 574
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 321 NRRKYFTNSYMYTJASICGEAYWGNYMFRGCSNVRPNI RVSKMSDGFTYMENSDDRRLYI 380
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 575 NSPYESTLD---THALSSQ-----PMLSISPASDNQLOSEIHDPGLNV 615
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 361 TKHD-QG---GWG-TLDEDPEGDOGHMR-----FIPL-RKGKTYV 414
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 616 PHYGOGLMTGMAWTOPEPASATITDPORANFRHTLLTLWLPAQVVPSPKHSRPLI 675
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
OY 415 SSKRMPNPMFMYESASGYIRSNEMNPPOGHWSIT 450
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
DB 676 ANTLMGNNMLATESLKN-----SALTLPFGCHPRPWGIT 707
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
RESULT 11
AAAG83208
ID AAG83208 standard; Protein: 977 AA.
AC AAC83208;
XX 05-SEP-2001 (first entry)
DT

Db 663 ANTLMGNNMLATESLKN-----SALTPSGHPFWGIT 694

RESULT 8

AAG83202
ID AAG83202 standard; Protein; 964 AA.
AC AAG83202;

DT 05-SEP-2001 (first entry)

DE Protein encoded by Chlamydia trachomatis pmpe gene.

XX Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; CapI; CT529; OMCB;
XX polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
OS Chlamydia trachomatis.

PN WO20014074-A2.

PD 07-JUN-2001.

PF 04-DEC-2000; 2000WO-US32919.

PR 03-DEC-1999; 99US-0454684.

PR 19-APR-2000; 2000US-0556877.

PR 20-JUN-2000; 2000US-0598419.

PA (CORI-) CORIXA CORP.

PI Probst P, Bhatia A, Skelky YAM, Fling SP, Scholler J;

DR WPI: 2001-374831/39.

XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease.

BS Claim 2; Page 191-193; 295pp; English.

XX The present sequence is provided in a specification relating to
XX compounds and methods for the treatment and diagnosis of chlamydial
XX infection. The compounds provided include polypeptides and fusion
XX proteins comprising immunogenic portions of Chlamydia antigens
XX and DNA sequences encoding such polypeptides. They are useful for
XX vaccinating against chlamydial infection, which causes pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease.

SO Sequence 964 AA;

Query Match Best Local Similarity 4.8%; Score 112.5; DB 22; Length 964;
Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

Db 21 GKRSPV-----NSILTKVEALKEAGSNEALEGKGEIOTPDPRVGATKILGSVG 76

Db 254 GNRGNIVYNNRCFRNVTASSEAS-----DGAIVTTRLDVTGNNGRIFFS-- 301

Qy 77 SALCKLNSGATKIIISGCLDIVAGIATFGPGVGMGIGAVAFSSILSLTGSAAKNSV 136

Db 302 -----DNTITNYGAI--YAPVITLVDSNPPTFFINNINANKG 336

Qy 137 AAVIDRALSKHREDAIORHA-----AGAKR-----DFAESSATIQV 172

Db 337 GATYIDGTNSKISA--DRAHITENENIVNVANSTSTSANPPRRNATIVASSSEHLL 395

Qy 173 MKQOS--NLTDSD--LSIIANVPYKFSNFISIGLESRIISGAAATTSIDSKRAVD----- 224

Db 396 GAGSSONLIFDPIEVSNAGVSV--SFNKEADDTGTVSPGATVNSADDFQRMLOTRTPAP 454

Qy 225 -----FILL--TCQLVVMRETL---LVDL---AIL--YKGAHEHVASAVENANRYNKE 268
Db 455 LTISSNGFLCTEDHQAOLTVNRFTQTGGVSLGNGAVLSCKYKNGTGD-----SASNASTTLKH 510
Qy 269 LAADTLDFLHKILIPBQALGAVYHPISASETSKAILNYT-----KFGVDVPRPI-----G 320
Db 511 IGLN-----LSSTLKSGAEL-----PLWVEPTNNNNNTADTAARFSLSDVLTSLIDYG 561
Qy 321 NRRYFTNSYMYVYSICSEAVYNGNMFGCSVNRPNIRVSKMSDGFYTMNSDRKLYT 380
Db 562 NSPYESTDL-----THALSSQ-----PMLTISEADNOIQSENIDFSGILNV 602
Qy 381 TKHD--QG---WGWG--TLDEDPDGDGHR-----FIPL--RNGRYNV 414
Db 603 PHIGWGLMTWGMVAKTQDDPPASATITTDQKRRFRTLLTWLPAGYTPSPRKHNSPLI 662
Qy 415 SSRKMPWPMYXMESSASAGTIRSWENNPQGHWSIT 450
Db 663 ANTLMGNNMLATESLKN-----SALTPSGHPFWGIT 694

RESULT 9

ABB94173
ID ABB94173 standard; Protein; 964 AA.
AC ABB94173;

DT 05-JUN-2002 (first entry)

DE Chlamydia protein sequence SEQ ID NO:177.

XX Chlamydia; vaccine; infection; Chlamydia; vaccine; detection; diagnosis;
XX antigen; antibacterial; immunostimulant; immune response;
XX Chlamydia-specific T-cell response.

OS Chlamydia sp.

PN WO200208267-A2.

PD 31-JAN-2002.

PF 20-JUL-2001; 2001WO-US23121.

PR 20-JUL-2000; 2000US-0620412.

PR 23-APR-2001; 2001US-0841132.

PA (CORI-) CORIXA CORP.

PI Fling SP, Skelky YAM, Probst P, Bhatia A;

DR WPI: 2002-179901/23.

XX Novel compositions comprising Chlamydia CapI protein and its use in the
XX treatment of Chlamydia infection.
XX Example 1; Page 223-225; 537pp; English.

XX The present invention describes compositions comprising a Chlamydia CapI
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment and/or expand T cells
XX infection. A192394 to A192709 and ABB94096 to ABB94374 represent
XX sequences used in the exemplification of the present invention.
SO Sequence 964 AA;


```

Db 1176 ISATNPEMN-PDITNOKASOVN--SAKSALNGDEK-----LAAAKQAKSDIGR 1222
OY 246 LYRKGAHEHVASAVENANRVNKELAADTLDFLHKLIPEQALIGAVYHPISASSETSALIN 305
Db 1223 LFDLNNAGQTSATAEVDAPN--LAAVTA-AKKNKATSLNTAMGNLKHAEKDNTRRSVN 1279
OY 306 YT 307
Db 1280 YT 1281

RESULT 5
AU37017
ID AU37017 standard; Protein; 5795 AA.
AC AU37017;
XX 14-FEB-2002 (first entry)
Df
Df Staphylococcus aureus cellular proliferation protein #1187.
Df
Df Antisense; prokaryotic cellular proliferation protein;
Df antibiotic; antibacterial; drug design.
Df
Df Staphylococcus aureus.
Df
Df WO200170955-A2.
Df
Df 27-SEP-2001.
Df
Df 21-MAR-2001; 2001WO-US09180.
Df
Df 21-MAR-2000; 2000US-191078P.
Df 23-MAY-2000; 2000US-206848P.
Df 26-MAY-2000; 2000US-207727P.
Df 23-OCT-2000; 2000US-242578P.
Df 27-NOV-2000; 2000US-253625P.
Df 22-DEC-2000; 2000US-257931P.
Df 16-FEB-2001; 2001US-269308P.
Df
Df (ELIT-) ELITRA PHARM INC.
Df
Df Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Df Yamamoto RT, Xu HH;
Df WPI: 2001-611495/70.
Df N-PSDB: AAS54876.
Df
Df New polynucleotides for the identification and development of
Df antibiotics, comprise sequences of antisense nucleic acids -
Df
Df Example 3; Seq ID No 12610; 511pp; English.
Df
Df The invention relates to antisense inhibitors of genes essential to
Df prokaryotic cellular proliferation, their use in identifying the
Df genes, their use in the discovery of novel antibiotics, the essential
Df genes themselves and the encoded proteins. The prokaryotes used are
Df Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
Df pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
Df invention is also useful for the identification of potential new targets
Df for antibiotic development. The antisense nucleic acids can also be used
Df and to obtain antibodies capable of binding to the expressed proteins.
Df The proteins can be used to screen compounds in rational drug discovery
Df programmes. The antisense nucleic acid sequence is also useful to screen
Df for homologous nucleic acids which are required for cell proliferation in
Df a wide variety of organisms. The present sequence represents an
Df essential prokaryotic cellular proliferation protein.
Df
Df Note: The sequence data for this patent did not form part
Df of the printed specification, but was obtained in electronic
Df format directly from WIPO at
Df ftp.wipo.int/pub/published_pct_sequences.

```

```

XX SQ Sequence 5795 AA;
Query Match 5.1%; Score 118.5; DB 22; Length 5795;
Best Local Similarity 20.9%; Pred. No. 5.1;
Matches 63; Conservative 59; Mismatches 121; Indels 59; Gaps 12;
OY 16 TSARKGRSDVNSLTFTKVTALKEAGSGNE--ALEGKGELQKRPDRV 65
Df 4486 TPADREKQTAIVDAVTAAKTILDKTAGSNDKAAVEQALQRYNTAKTALNGDA-----RL 4540
OY 66 GOATKILGSVSGALGKLSNGDATKIIISGCLDIAGIATFGGPGVGMIGAVASFSSILS 125
Df 4541 NEAKNTAKQOQLATMSHLNNAOKANLT--QIERG--TTVAVQGI----- 4581
OY 126 LFTGSSAKNSVAIVDRALSKHRDIAIQRHAAKARDPRESSAFIQVMKQSNLTDSLS 185
Df 4582 -----QANAGTLDQAMNOLROSIAKSDATKSESDYODANADLQNNAYNHA--VSDAEGT 4632
OY 186 IIAANVPYKFSNFIQLESRTSGAATTSLSDAKRAVDILLYQQLVVMRETLVLDLAI 245
Df 4633 ISATNPEMN-PDITNOKASOVN--SAKSALNGDEK-----LAAAKQAKTDIGH 4679
OY 246 LYRKGAHEHVASAVENANRVNKELAADTLDFLHKLIPEQALIGAVYHPISASSETSALIN 305
Df 4680 LSDLNNAGQTSATAEVDAPN--LAAVTA-AKKNKATSLNTAMGNLKHAEKDNTRRSVN 4736
OY 306 YT 307
Df 4737 YT 4738

```

```

RESULT 6
AU37403
ID AU37403 standard; Protein; 6281 AA.
AC AU37403;
XX 14-FEB-2002 (first entry)
Df
Df Staphylococcus aureus cellular proliferation protein #1573.
Df
Df Antisense; prokaryotic cellular proliferation protein;
Df antibiotic; antibacterial; drug design.
Df
Df Staphylococcus aureus.
Df
Df WO200170955-A2.
Df
Df 27-SEP-2001.
Df
Df 21-MAR-2001; 2001WO-US09180.
Df
Df 21-MAR-2000; 2000US-191078P.
Df 23-MAY-2000; 2000US-206848P.
Df 26-MAY-2000; 2000US-207727P.
Df 23-OCT-2000; 2000US-242578P.
Df 27-NOV-2000; 2000US-253625P.
Df 22-DEC-2000; 2000US-257931P.
Df 16-FEB-2001; 2001US-269308P.
Df
Df (ELIT-) ELITRA PHARM INC.
Df
Df Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Df Yamamoto RT, Xu HH;
Df WPI: 2001-611495/70.
Df N-PSDB: AAS55262.
Df
Df New polynucleotides for the identification and development of
Df antibiotics, comprise sequences of antisense nucleic acids -
Df
Df Example 3; Seq ID No 12996; 511pp; English.

```

XX OS Chiroptalmus quadrigatus.
XX PN JP2002125676-A.
XX PD 08-MAY-2002.
XX PF 24-OCT-2000; 2000JP-0323506.
XX PR 24-OCT-2000; 2000JP-0323506.
XX PA (SUNR) SUNTORY LTD.
XX DR WPI: 2002-483935/52.
XX DR N-PSDB: AAL47248.
XX PT New hemolytically active protein useful as intermediates for
XX PT agricultural chemicals
XX PS Claim 2; Page 8-9; 13pp; Japanese.
XX CC The present invention relates to proteins from the sea wasp which have
XX CC haemolytic activity. The proteins can be used for the preparation of
XX CC reagents, agricultural chemicals and drugs. The present sequence is a
XX CC polypeptide of the invention.
XX SQ Sequence 462 AA;
Query Match 18.2%; Score 423.5; DB 23; Length 462;
Best Local Similarity 26.3%; Pred. No. 1.9e-28;
Matches 120; Conservative 108; Mismatches 171; Indels 57; Gaps 17;
QY 24 SDVNSLTLYETALKEASGSEALALEGLKEIQTKEPRVQATKILGSVGS---ALG 80
DB 26 SGLASLAKIDA--KRPSG--KQLEFDYANNQKQIEEFKFSNDDERAKMGAGISSTAVG 81
QY 81 KLNSGDATKIISGCLDIAGIATTFGGPYGMCIGAVASFVSSIIISLFTGSAAKNSVAVI 140
DB 82 KFOSGDPAKTASGCLDIIVGIVSVLKD--FAKSPISITLSMVGILSGTAEESVGSV 139
QY 141 DRALSKHRDEAIORHAAGAKRDEAFESSAFID-VAKQOONLTDLSIIAANVPYKFSNF 199
DB 140 KKVVOESQDDELQALYGVKREYAVSKAFIDGVNENSDLSPTVSAAGANVPYQGVRF 199
QY 200 IGLAESISGAATTSISDAKRAVDILYQVVMRETLVLD-ALTYRKGNAEHVASA 258
DB 200 IAAVVOIRIKNNKPRTE-SEIKRVISMELFTDLCSLRDLILDLVQVAVFGHSPNIAAG 258
QY 259 VENANRVNKEIADTLDFLHKLIEQALIGAVY-HPISASETSKAILNYTYFGVDPVPR 317
DB 259 IKEYSNIGREYKRVFEDLKTNDKETYLLFSLYPRRNOSOKIF---KFF---DLMK 312
QY 318 PIGNRRYK-----FTNSYWNTRYSSICEAYMGNMFRGCSNVNPNIRVSKMSDGEYT 369
DB 313 VKYDRLKQDLDTGIVFSSLSLHMPNYFLCSSK---DYALLICTKPYG-SLRDLKNDGRYS 368
QY 370 MENSDDR-----KLYTHDQGWGWTDEDPDGGHRRFPLRKG---YAVSSKR 418
DB 369 IKTKQSNPKYCHRYEYILFTFHDRN-----DDELEKNEFVPVKIGEKRIYLLSSKA 418
QY 419 WPNMFMTESASGYIRSMENN-----GPOGHMSI 449
DB 419 SPNFATVVPKTAAGDLEFFVDGIPSOGLGNOGYFTL 454
RESULT 4
AAU34339
ID AAU34339 standard; Protein; 2434 AA.
XX AC AAU34339;
XX DT 14-FEB-2002 (first entry)
XX

DE Staphylococcus aureus cellular proliferation protein #615.
XX KM Antisense; prokaryotic cellular proliferation protein;
XX KM antibiotic; antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELITR) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI: 2001-611495/70.
XX DR N-PSDB: AAS52198.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 5835; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhii, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pcl_sequences.
XX SQ Sequence 2434 AA;
Query Match 5.1%; Score 118.5; DB 22; Length 2434;
Best Local Similarity 21.2%; Pred. No. 1.4;
Matches 64; Conservative 58; Mismatches 121; Indels 59; Gaps 12;
QY 16 TSAHGRSDVNSLTLYETALKEASGSE--FALE-----ALGLKEIQTKEPRV 65
DB 1029 TDADREKOTAVNDAVTAKTILDTAASNDKRAVEBALORVNTAKTALMGD-----ERL 1083
QY 66 GQATKILGSVSGALKINSGDATKIISGCLDIYAGIATTFGGPYGMCIGAVASFVSSILS 125
DB 1084 NEAKNTAKQOVATMSHLTDQOKANLTS---QIESG--TVVAGVGGI----- 1124
QY 126 LFTGSSAKNSYAAVYIDRALSKHRDEAIORHAAGAKRDEAFESSAFIDVAKQOONLTDSDLS 185
DB 1125 -----QANAGTLDDAMNOLROSIAKSDTKSSSEYODANMDLQANAYNDA-VTNAEGI 1175
QY 186 IIAANVPYKFSNFIGLESRIISGAATTSISDAKRAVDILYQVVMRETLVLDLAI 245

PT mechanism
 XX
 PS Claim 4; Page 26-27; 32pp; Japanese.
 XX
 CC This invention describes a novel protein which has hemolytic activity,
 CC blood platelet agglutination activity and a molecular weight of about
 CC 50,000 Da (by SDS-PAGE). The protein, homologs and analogs of the
 CC invention can be used as a novel approach to developing drugs useful for
 CC the treatment of jelly fish stings, pharmaceuticals with blood platelet
 CC agglutination activity, pesticides by use of the hemolytic activity, and
 CC in the study of the hemolytic mechanism. This sequence represents the
 CC hemolytic protein described in the invention
 XX
 SQ Sequence 450 AA;
 Query Match 100.0%; Score 2322; DB 20; Length 450;
 Best Local Similarity 100.0%; Pred. No. 2,5e-195;
 Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MILKHPMLFIVLAITSAGKGRSDVNSLITFETALKRAGSNEALALGELGEIOT 60
 DB 1 MILKHPMLFIVLAITSAGKGRSDVNSLITFETALKRAGSNEALALGELGEIOT 60
 QY 61 KPDVGOATKILGVSALGKLSGDAATKISGCLDIYAGIATFEGGPGVGMIGAVASPV 120
 DB 61 KPDVGOATKILGVSALGKLSGDAATKISGCLDIYAGIATFEGGPGVGMIGAVASPV 120
 QY 121 SSILSFTSSAKNSVAVIDRALSKRDEAIORHAAGKRDFAESSAFIOVMKQOQSNLT 180
 DB 121 SSILSFTSSAKNSVAVIDRALSKRDEAIORHAAGKRDFAESSAFIOVMKQOQSNLT 180
 QY 181 DSDLSITIANVPYKESNFTIGLESRIISGAAATSLSDAKRAVDFTLLCOLVVMRETL 240
 DB 181 DSDLSITIANVPYKESNFTIGLESRIISGAAATSLSDAKRAVDFTLLCOLVVMRETL 240
 QY 241 VLALITIRKGNAEHVASAVENANRVKELADTLDLHLKLIPEALIGAVYHPISASETS 300
 DB 241 VLALITIRKGNAEHVASAVENANRVKELADTLDLHLKLIPEALIGAVYHPISASETS 300
 QY 301 KAILNTKTFGVDPVPRPIGNRRYKFTNSYNTYSICSEAYMGNYMFRGCSNVRNPNIY 360
 DB 301 KAILNTKTFGVDPVPRPIGNRRYKFTNSYNTYSICSEAYMGNYMFRGCSNVRNPNIY 360
 QY 361 SKMSDFEYTMENSDRRKLTITTKHDOGMGWGTLDLDEPDGOGHMFIFLRHGKYVSSKRP 420
 DB 361 SKMSDFEYTMENSDRRKLTITTKHDOGMGWGTLDLDEPDGOGHMFIFLRHGKYVSSKRP 420
 QY 421 NMFMWESSASGYIRSMENNPQGHWSIT 450
 DB 421 NMFMWESSASGYIRSMENNPQGHWSIT 450
 RESULT 2
 AAG65633
 ID AAG65633 standard; protein; 463 AA.
 XX
 AC AAG65633;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE C. alata hemolytic protein.
 KW Hemolytic protein; nematocyst; sting; drug; blood platelet; agrochemical.
 XX
 OS Carybdea alata.
 XX
 PN WO200170799-A1.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-JP02209.
 XX
 PR 21-MAR-2000; 2000JP-0078967.

XX
 PA (SUNR) SUNTORY LTD.
 XX
 XI Nagai H, Kuroda K, Nakajima T;
 XX
 PI WPI: 2001-611481/70.
 DR N-PSDB: AA166621.
 XX
 CC New polypeptide for treating stings, and for the development of
 CC pharmaceuticals and agrochemicals, and studying hemolysis, comprises
 CC the Carybdea alata-originated protein
 CC
 CC Claim 4; Page 22-23; 28pp; Japanese.
 PS
 XX
 CC The invention provides a polypeptide from nematocyst of C. alata. The
 CC protein comprises hemolytic activity, and has a molecular weight of about
 CC 50,000 Daltons as determined by SDS-polyacrylamide gel electrophoresis
 CC (SDS PAGE). The protein is useful for treating stings, and also in the
 CC development of drugs with blood platelet aggregating effects, for the
 CC development of agrochemicals with hemolytic effects, and as reagents for
 CC studying hemolytic mechanisms. The present sequence represents the
 CC C. alata protein having hemolytic activity.
 CC
 SQ Sequence 463 AA;
 Query Match 42.7%; Score 992.5; DB 22; Length 463;
 Best Local Similarity 43.5%; Pred. No. 1.8e-78;
 Matches 197; Conservative 95; Mismatches 148; Indels 13; Gaps 7;
 QY 9 LFIIVLAITSAGK-----KRSVNSLITFETALKRAGS--GSEAL--EALGELGEIOT 60
 DB 11 LFIIVLAITSAGKRLSRYSRAADAVSTDIDGILGOLNDLGTFRKLEALOGVOEAVK 70
 QY 61 KP-DVGOATKILGVSALGKLSGDAATKISGCLDIYAGIATFEGGPGVGMIGAVASPV 119
 DB 71 EATYIAKSTIVGSGSLKFKSGDPDVASGCLDIYAGIATFEGGPGVGMIGAVASPV 130
 QY 120 VSSILSFTSSAKNSVAVIDRALSKRDEAIORHAAGKRDFAESSAFIOVMKQOQSNLT 179
 DB 131 ISSILSFTSSAKNSVAVIDRALSKRDEAIORHAAGKRDFAESSAFIOVMKQOQSNLT 190
 QY 180 TSDLSITIANVPYKESNFTIGLESRIISGAAATSLSDAKRAVDFTLLCOLVVMRETL 239
 DB 191 TEVHLDSDVDAVRAFTMLGVLESIRNGSVSDNNAMRTINFTFLYDLSVNRRL 250
 QY 240 LVDLALYK--GNAEHVASAVENANRVKELADTLDLHLKLIPEALIGAVYHPISAS 297
 DB 251 LQVILITIRKAGAYDELALSLISDQKKEATREIVTLHOMETKYSICGSYIYFIDS 310
 QY 298 ETSKALNTKTFGVDPVPR-PIGNRRYKFTNSYNTYSICSEAYMGNYMFRGCSNVRNP 356
 DB 311 KALIGILKLTKEFGVDPVPRPIGNRRYKFTNSYNTYSICSEAYMGNYMFRGCSNVRNP 370
 QY 357 NIRVSKMSDFEYTMENSDRRKLTITTKHDOGMGWGTLDLDEPDGOGHMFIFLRHGKYVSS 416
 DB 371 GIRIKKLENGYHTI--TLRSKAMVYTKHAOGMGWGTLDLDEPDGOGYFTFLPLNGFYMAST 429
 QY 417 KRMPNFMWESSASGYIRSMENNPQGHWSI 449
 DB 430 KRMPDYFVYMESSAHGYIRSMHYNPDQGMKI 462
 RESULT 3
 AAO17909
 ID AAO17909 standard; peptide; 462 AA.
 XX
 AC AAO17909;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Sea wasp hemolytic protein SEQ ID NO: 3.
 XX
 KW Sea wasp; haemolytic; agricultural chemical; drug.

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 11:14:48 ; Search time 77 Seconds

(without alignments)
778,737 Million cell updates/sec

Title: US-09-647-522-5

2322

Sequence: 1 MILKHPMLRIVLAITSARK.....SGYRSMENNGPGHWSIT 450

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2322	100.0	450	AA133651	C. rastoni hemolytic
2	992.5	42.7	463	AA65633	C. alata hemolytic
3	423.5	18.2	462	AA017909	Sea wasp hemolytic
4	118.5	5.1	2434	AAU34339	Staphylococcus aur
5	118.5	5.1	5795	AAU37017	Staphylococcus aur
6	118.5	5.1	6281	AAU37403	Staphylococcus aur
7	112.5	4.8	964	AA133634	C. trachomatis pmp
8	112.5	4.8	964	AA683202	Protein encoded by
9	112.5	4.8	964	AB894173	Chlamydia protein
10	112.5	4.8	977	AA133640	C. trachomatis pmp

11	112.5	4.8	977	22	AA683208
12	112.5	4.8	977	23	AB894179
13	110	4.7	1066	21	AA649832
14	110	4.7	1073	21	AA649831
15	110	4.7	1086	21	AA649830
16	106	4.6	1978	20	AA649830
17	105.5	4.5	765	23	ABP25599
18	105.5	4.5	956	23	AAE33466
19	105.5	4.5	965	23	AAE33453
20	105.5	4.5	1844	18	AAW18302
21	105.5	4.5	1844	19	AAW56558
22	105.5	4.5	2504	18	AAW17871
23	105.5	4.5	2504	19	AAW56557
24	105.5	4.5	2504	22	AAW72610
25	105.5	4.5	2505	22	AAW72612
26	104.5	4.5	837	22	AAU34387
27	104.5	4.5	875	22	AAU37487
28	104.5	4.5	989	20	AAU37242
29	104.5	4.5	1788	22	AB862995
30	104.5	4.5	2086	22	AAU34143
31	103.5	4.5	349	22	AAW49636
32	103.5	4.5	565	22	AAW49642
33	101	4.3	1250	22	ABW12254
34	100.5	4.3	847	22	AAW40081
35	100.5	4.3	1645	18	AAW13502
36	100	4.3	927	20	AAW35448
37	100	4.3	927	23	ABW49720
38	100	4.3	3158	22	AAU37018
39	99	4.3	885	21	AAW32050
40	99	4.3	931	21	AAW32049
41	99	4.3	956	21	AAW32048
42	97.5	4.2	1013	19	AAW56013
43	97.5	4.2	1149	21	AAW16683
44	97.5	4.2	1445	12	AAW12108
45	97.5	4.2	2181	22	ABW05687

ALIGNMENTS

RESULT 1
AA133651 standard; Protein: 450 AA.

AC AA133651;

DT 06-JAN-2000 (first entry)

DE C. rastoni hemolytic protein.

KW Hemolytic protein; blood platelet agglutination; drug development;
 RW treatment; sting; jellyfish; pharmaceutical; pesticide.

OS Carybdea rastoni.

PN W09950294-A1.

PD 07-OCT-1999.

PF 30-MAR-1999; 99WO-JP01607.

PR 01-APR-1998; 98JP-0088569.

PA (SUNR) SUNTORY LTD.

PI Nagai H, Nakajima T;

DR WPI: 1999-580740/49.

XX N-PSDB: AA23610.

XX Protein with hemolytic activity, useful for drugs treating jelly fish
 XX stings, pharmaceuticals with blood platelet agglutination activity,
 XX pesticides by use of the hemolytic activity, and study of the hemolytic